Good Firms, Worker Flows and Local Productivity

Michel Serafinelli
University of Toronto*

Abstract

Localized knowledge spillovers are a common explanation for the productivity advantages of agglomeration. Nevertheless if information can easily flow out of firms, the question of why the effects of spillovers are localized must be clarified. If knowledge is embedded in workers and diffuses when workers move between firms, the strong localized aspect of knowledge spillovers may arise at least in part from the propensity of workers to change jobs within the same local labor market. In this paper I present direct evidence on the role of firm-to-firm labor mobility in enhancing the productivity of firms located near highly productive firms. Using Social Security earnings records for workers matched to detailed financial data for their employers in Veneto, a region of Italy with many successful industry clusters, I first identify a set of highly productive firms. I then show that hiring a worker with experience at highly productive firms significantly increases the productivity of other (non-highly productive) firms. I do so using different techniques, including an instrumental variable strategy which exploits downsizing events at highly productive firms. Back-of-the-envelope calculations suggest that worker flows can explain around 10% of the productivity gains experienced by other firms when new highly productive firms are added to a local labor market.

JEL: R10; D24; J31; J60. Keywords: productivity, agglomeration advantages, linked employer-employee data, labor mobility.

*Contact Information: Department of Economics, University of Toronto, 150 St George St Toronto, ON M5S 3G7, Canada, michel.serafinelli@utoronto.ca, +1 416-978-4622. I am very grateful to David Card, Enrico Moretti, Patrick Kline and Bronwyn Hall for invaluable guidance. I also thank Miguel Almunia, Vladimir Asriyan, Audinga Baltrunaite, Gustavo Bobonis, Alex Bryson, Pamela Campa, Jeff Chan, Kunal Dasgupta, Francesco Devicienti, Sabrina Di Addario, Frederico Finan, Lee Fleming, Albrecht Glitz, Yuriy Gorodnichenko, Tadeja Gracner, Jonas Hjort, Mitchell Hoffman, Marco Leonardi, Shimeng Liu, Lisa Lynch, Agata Maida, Robert McMillan, Peter Morrow, Michele Pellizzari, Marco Percoco, Giuseppe Ragusa, Steven Raphael, Ana Rocca, Andrés Rodríguez-Clare, Jesse Rothstein, Fabio Schiantarelli, Fabiano Schivardi, Shalini Sharma, Aloysius Siow, Christopher Surfield, Grigorios Spanos, Matthew Turner, Victoria Vanasco, Yves Zenou and seminar participants at NBER Summer Institute, Wharton, UPF, Warwick, PSE, SOLE, Berkeley, U of Calgary, U of Cambridge, Stockholm U, UAB, U of Alicante, Collegio Carlo Alberto, U of Mannheim, fRDB, EIEF, UTDT, Urban Economic Association Meeting, U of Toronto for comments. I acknowledge financial support from the Berkeley CEG. Thanks are due also to Federico Callegari and to Veneto firms' employees and officials of employers' associations who have generously given their time in interviews, and to Jeff Chan, Grigorios Spanos, Jessica Burley, Ishita Arora, Olivia Casey and Nitin Kohli for excellent research assistance. I am indebted to Giuseppe Tattara for making available the VWH Data.

1 Introduction

A prominent feature of the economic landscape in the most developed countries is the tendency for firms to locate near other firms producing similar products or services. In the United States, for example, biopharmaceutical firms are clustered in New York and Chicago and a sizeable share of the elevator and escalator industry is concentrated in the area around Bloomington, Indiana. In addition, the growth and diffusion of multinational corporations has led to the recent appearance of important industrial clusters in several emerging economies. Firms that originally agglomerated in Silicon Valley and Detroit now have subsidiaries clustered in Bangalore and Slovakia (Alfaro and Chen, 2014).

Researchers have long speculated that firms in industrial concentrations may benefit from agglomeration economies, and a growing body of work has been devoted to studying the importance of these economies. Despite the difficulties involved in estimating agglomeration effects, a consensus has emerged from the literature that significant productivity advantages of agglomeration exist for many industries (Rosenthal and Strange, 2003; Henderson, 2003; Cingano and Schivardi, 2004; Ellison, Glaeser and Kerr, 2010; Greenstone, Hornbeck and Moretti, 2010; Combes et al., 2012). Localized knowledge spillovers are a common explanation for the productivity advantages of agglomeration.[1] Nevertheless, as pointed out by Combes and Duranton (2006), if information can easily flow out of firms, the question of why the effects of spillovers are localized must be clarified.

This paper directly examines the role of labor mobility as a mechanism for the transfer of efficiency-enhancing knowledge and evaluates the extent to which labor mobility can explain the productivity advantages of firms located near other highly productive firms. The underlying idea is that knowledge is embedded in workers and diffuses when workers move between firms. The strong localized aspect of knowledge spillovers discussed in the agglomeration literature may thus at least in part arise from the propensity of workers to change jobs within the same local labor market.

In order to empirically assess the importance of labor-market based knowledge spillovers, I use a unique dataset that combines Social Security earnings records and detailed financial information for firms in Veneto, a region of Italy with many successful industry clusters. I find that hiring a worker with experience at a highly productive firm increases the productivity of other (non-highly productive) firms, and that worker flows explain a significant portion of the productivity gains experienced by other firms when highly productive firms in the same industry are added to a local labor market.

I begin by presenting a simple conceptual framework where some firms are more pro-

[1] The availability of specialized intermediate inputs, the sharing of a labor pool, and better matching have also garnered attention in the literature's attempt to explain agglomeration economies.

ductive because they have some superior knowledge. Employees at these firms passively acquire some proportion of the firm's internal knowledge. For simplicity, I refer to these as "knowledgeable" workers. Other firms can gain access to the superior knowledge by hiring these workers. Empirically, I identify potential high-productivity firms as those that pay a relatively high firm-specific wage premium.[2] I show that these high-wage-firms (HWFs) have significantly higher total factor productivity (TFP) and value added than other firms in my sample, suggesting the presence of a firm-specific advantage and thus a point of origin for the transfer of knowledge. Next, I evaluate the extent to which non HWFs benefit from hiring knowledgeable workers by studying the effect on productivity associated with hiring workers with recent experience at HWFs.

My investigation is subject to identification concerns. An obvious one is that firms that hire workers with recent HWF experience are different than those that do not, and that this underlying difference – rather than knowledge transfers – account for the measured productivity effects. In particular, productivity shocks that are correlated with the propensity to hire knowledgeable workers may give rise to an upward bias in the impact of knowledgeable workers. In order to address this potential endogeneity issue, I use control function methods in the spirit of the productivity literature (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg, Caves, and Frazer, 2008).

It is also possible that knowledgeable workers are attracted to join firms that are on the rise, rather than knowledgeable workers moving to firms and causing the increase in productivity. To explore this possibility, I instrument for the number of knowledgeable workers in a non HWF with the number of local good firms in the same industry that downsized in the previous period.[3] Indeed, following a downsizing event at a HWF, it is more likely that a knowledgeable worker applies for job at local non HWFs because s/he is unemployed and does not want to relocate far away, and less likely that s/he does so because some particular non HWF offers better prospects than the HWF at which the worker is

[2] Results are similar when using alternative groupings of firms based on output (controlling for inputs) and value added. The definition of potential high-productivity firms as high-wage firms (HWFs) employed in my baseline analysis is consistent with many recent models of frictional labor markets (e.g., Christensen et al., 2005), in which higher-productivity firms pay higher wages for equivalent workers. There are three reasons why for the baseline results I define the good firms as high-wage-firms and detect them using Social Security data rather than define the good firms directly as the highly productive ones and detect them using balance sheet data. First, the availability of worker-level Social Security data allows the introduction of measured individual characteristics and worker effects, something impossible to capture with firm level data from balance sheets. Second, Social Security data are available for a longer period of time than the balance sheets, and therefore increase the precision of the categorization of firms into good and non-good groups. Third, since Social Security records are administrative data, measurement error is lower than in balance sheets.

[3] My instrument is an external 'z'-variable used in a production function framework. As pointed out in the survey Eberhardt and Helmers (2010), to date only past values of the regressors themselves or input prices have been used for instrumentation in the production function literature.

employed[4].

A final potential threat to identification is the fact that I do not observe labor quality. In order to investigate this issue, I obtain a proxy for worker ability by procuring estimates of worker fixed effects from wage equations where both firm and worker effects can be identified.

I conclude that hiring a worker with HWF experience significantly increases the productivity of other (non HWF) firms. Greater productivity gains are observed in firms hiring workers in higher-skilled occupations. The productivity effect of knowledgeable workers is not associated with recently hired workers in general; I do not find a similar productivity effect for recently hired workers with experience at firms which have lower productivity than the receiving firm.[5] While I cannot completely exclude the possibility that at least some of the estimated effect reflects better worker-firm matching, or switchers being more productive than stayers in general (i.e. regardless of the previous employment history), this evidence lends credibility to the knowledge transfer hypothesis. Overall, while none of the individual empirical tests can fully rule out alternative explanations, they offer evidence that is consistent with a causal interpretation of the observed associations between worker flows and productivity.[6]

In the last part of the paper, I evaluate the extent to which labor mobility can explain the productivity advantages of firms located near other highly productive firms. I relate my findings on the effect of firm-to-firm labor mobility to the existing evidence on the productivity advantages of agglomeration, focusing in particular on the study performed in Greenstone, Hornbeck and Moretti (2010, henceforth GHM). The authors 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 increases significantly relative to the TFP of incumbent plants in counties that survived a long selection process but narrowly lost the competition. The observed effect on TFP is larger if incumbent plants are in the same industry as the large plant, and increases over time. These two facts are consistent with the presence of intellectual externalities that are embodied in workers who move from firm to firm. However, data limitations prevent GHM from drawing definitive conclusions regarding the driving mechanism. I evaluate the extent to which worker flows explain empirical evidence on the productivity advantages of agglomeration, by predicting, within the worker mobility framework described above, the change in local productivity following an

[4] As an alternative approach to address this issue, I adapt the control function methods to proxy for future productivity shocks.

[5] Also, if a positive and significant relationship between labor mobility and productivity was driven by higher ability of workers moving from HWFs, then the coefficient on recent hires from firms with productivity lower than the receiving one would be negative and significant, since hiring labor from non-HWFs would cause a decline in the firm's average worker quality and therefore deteriorate its productivity. I find that such coefficient is small and positive.

[6] I also show that the results are not driven by time-invariant unobservables such as managerial talent.

event analogous to that studied by GHM. The change in productivity predicted within this framework equals around 10 percent of the overall local productivity change observed after the event.

Finally, I show that the local productivity effect attributed to good firms does not appear to be associated with an increase in the size of the labor market in general: large productivity gains linked to changes in the number of firms seem to be realized only when the new entrants are good firms. Although I am not able to entirely discard the chance that at least part of the estimated impact reflects better worker-firm matching arising from a thicker labor market, this finding supports the hypothesis of labor-market based knowledge spillovers.

The remainder of this paper is structured as follows. In Section 2, I relate my research to the existing literature. Section 3 presents a conceptual framework that guides the empirical exercise and discusses the econometric strategy. In Section 4, I describe my data and I present descriptive results. The main regression results, in addition to various extensions and robustness checks are presented in Section 5 and 6. Section 7 concludes the paper.

2 Relation to Previous Research

This paper adds to the growing literature on productivity advantages through agglomeration, a literature critically surveyed in Duranton and Puga (2004), Rosenthal and Strange (2004) and Moretti (2011). The research relating most closely to this paper is the body of work on micro-foundations for agglomeration advantages based on knowledge spillovers. In Combes and Duranton (2006)'s theoretical analysis, 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 key employees to competitors along with higher wage bills to retain other key employees).[7] In a case study of the British Motor Valley, Henry and Pinch (2000) conclude that

> as personnel move, they bring with them knowledge and ideas about how things are done in other firms helping to raise the knowledge throughout the industry...The crucial point is that whilst this process may not change the pecking order within the industry, this 'churning' of personnel raises the knowledge base of the industry *as a whole within the region*. The knowledge community is continually reinvigorated and, synonymous with this, so is production within Motor Sport Valley

[7] The study of R&D spillover effects by Bloom, Schankerman and Van Reenen (2013) points out the presence of two countervailing effects: positive technological spillovers and negative business-stealing effects on the product market. The authors provide evidence that although both types of effects operate, technological spillovers quantitatively dominate.

In a similar vein, Saxenian (1994) maintains that the geographic proximity of high-tech firms in Silicon Valley is associated with a more efficient flow of new ideas. I contribute to the literature on micro-foundations for agglomeration advantages by showing direct evidence of productivity gains through worker flows. My results are consistent with the findings by Henry and Pinch (2000).

Some research beyond the agglomeration literature has also emphasized the fact that firm-to-firm labor mobility may enhance the productivity of firms. Dasgupta (2012) studies a dynamic general equilibrium model with mobility of workers among countries, in which the long-term dynamic learning process plays a crucial role. Workers in the model learn from their managers and knowledge diffusion takes place through labor flows. Other theoretical contributions are Cooper (2001), Markusen (2001), Glass and Saggi (2002) and Fosfuri, Motta and Rønde (2001). For what concerns the empirical literature using linked worker-firm data, Poole (2013) finds a positive effect on wages paid in domestic firms in Brazil of the share of new workers previously employed by foreign-owned firms. Balsvik (2011) offers a detailed account of productivity gains linked to worker flows from foreign multinational to domestic firms in Norway. Parrotta and Pozzoli (2012) show evidence for Denmark of a positive impact of the recruitment of knowledge carriers - technicians and highly educated workers recruited from a donor firm - on a firm's value added. Finally, Stoyanov and Zubanov (2012) show that Danish firms that hired workers from more productive firms experience productivity gains. My findings are consistent with those of these four recent empirical contributions. Unlike the above authors, who focus exclusively on the role of labor mobility for firm-level productivity, I seek to shed light on a broader question: the extent to which labor mobility can explain evidence on the productivity advantages through agglomeration. 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 of Italy where, as in the Silicon Valley, networks of specialized firms, frequently organized in districts, have been effective in promoting and adapting to technological change during the last three decades. This so called "Third Italy" region has received a good deal of attention by researchers, in the United States as well as in Europe (Brusco, 1983; Piore and Sabel, 1984; Trigilia, 1990; Piore, 2009). For what concerns the investigation on the firm-level productivity effect of hiring, my contribution with respect to Poole (2013), Balsvik (2011), Parrotta and Pozzoli (2012) and Stoyanov and Zubanov (2012) is to use a variety of approaches (IV, control function methods, placebo tests) to empirically tease out the different micro-level stories.

3 Framework and Empirical Strategy

3.1 Conceptual Framework

Assume there exists a finite number of locations, each constituting a separate local labor market. To fix ideas, assume that these labor markets are completely segmented with workers being immobile between them. There exists a finite collection $\mathcal{J} = \{\mathcal{J}_0, \mathcal{J}_1\}$ of firms consisting of the set $\mathcal{J}_1$ *good* firms, which are more productive because they have some superior knowledge and set $\mathcal{J}_0$ other firms which have no access to the superior knowledge. The superior knowledge is exogenously given and could include information about export markets, physical capital, process innovations, new managerial techniques, new organizational forms and intermediate inputs. Workers employed by good firms acquire some proportion of the firms' internal knowledge. For simplicity, I assume that this acquisition of internal knowledge takes place immediately after the workers join the good firm. Workers are *knowledgeable* if they have knowledge of the relevant information and *unknowledgeable* otherwise. All workers employed by good firms, then, are knowledgeable. Additionally, some proportion of this knowledge can be transferred to a $j \in \mathcal{J}_0$ firm if the workers switch employers.[8] The production function of firm $j \in \mathcal{J}_0$ is

$$Y_j = F(L_j, K_j, M_j) = A_j[{L_j}^{\alpha} K_j^{\gamma} M_j^{\lambda}]^{\delta} \tag{1}$$

where $L_j = H_j + N_j$, i.e. the sum of knowledgeable workers (H_j, who moved at some point from a good firm to a non-good firm) and unknowledgeable workers (N_j); K_j is total capital inputs, M_j is material inputs, and $\delta < 1$ represents an element of diminishing returns to scale, or to "span of control" in the managerial technology (Lucas, 1978).[9] I allow for knowledge transfer by letting productivity depend on H_j[10]:

$$A_j = D_j e^{\beta_H H_j} \tag{2}$$

3.2 Empirical Strategy

I obtain the regression equation that forms the basis of my empirical analysis, by combining equation (1) and (2) , and taking logs:

$$\ln(Y_{jslt}) = \beta_L \ln(L_{jslt}) + \beta_K \ln(K_{jslt}) + \beta_M \ln(M_{jslt}) + \beta_H H_{jslt} + \beta_0 + \zeta_{jslt} \tag{3}$$

[8] I assume that this type of knowledge cannot all be patented and that exclusive labor contracts are not available.

[9] This is in line with the large presence, that I document below, of small and medium size firms in the non-HWF sample.

[10] See Section A.I for a discussion of the firm optimization problem.

The dependent variable in much of my analysis is the real value of total firm production; s denotes industry, l denotes locality and t denotes year.[11] The variable of interest, H is constructed from head counts in the matched employer-employee data.[12] I define a worker as having recent HWF-experience in year t, if he or she is observed working in a HWF for one or more of the years $t-3$ to $t-1$. If a worker is hired at time $t-g$, and has experience at a HWF between $t-g$ and $t-3$, she contributes to H count from year $t-g$ until t. [13]

The structure of regression equation (3) is in line with that in GHM, who also regress firm-level output on inputs (and let productivity depend on the presence of large plants that generated bidding from local governments). The estimation of such productivity specification on balance sheet data allows me in Section 6 to relate directly my findings on the effect of firm-to-firm labor mobility to the evidence in GHM.

In equation (3) the term $\ln(D_j)$ is decomposed into two elements, β_0 and ζ_{jslt}. The constant β_0 denotes mean efficiency across all firms in $\mathcal{J}_0$ that is due to factors others than H. The time-variant ζ_{jslt} represents deviations from this mean efficiency level and captures (a) unobserved factors affecting firm output, (b) measurement error in inputs and output, and (c) random noise. Estimating the effect of recruiting a knowledgeable worker on a firm's productivity is difficult in the presence of unobservable productivity shocks (contemporaneous or future) and unobserved labor quality. I turn now to describing what type of biases these *time-varying* unobservables may introduce and how I deal with them in the empirical work.

In Section 5.3 I discuss estimates using the within-transformation, to address the possibility that the estimated productivity gains are due to *time-invariant* unobservables such as managerial talent.

3.2.1 Unobserved Productivity shocks

I express the deviations from mean firm efficiency not resulting from knowledge transfer, ζ_{jslt}, as

$$\zeta_{jslt} = \omega^*_{jslt} + \nu_{jslt} = \omega_{jslt} + \mu_{st} + \varpi_{lt} + \nu_{jslt} \tag{4}$$

[11]Subsection 5.3 reports estimation results for alternative specifications (in terms of functional forms and measures of productivity). Notice that $\beta_L = \delta\alpha, \beta_K = \delta\gamma, \beta_M = \delta\lambda$.

[12]In Section 5.4 I also employ an alternative, continuous, measure of the receiving firm's exposure to knowledge, which exploits the productivity differences between sending and receiving firms (in the spirit of Stoyanov and Zubanov (2012)), thus extending the baseline analysis which works with a dummy indicating experience at a HWF. Further, I present estimates when I lag the number of workers with HWF experience.

[13]It may be instructive to consider a practical example. Consider a worker who separates from a HWF in 1994 and joins non-HWF j in 1995. Provided that the worker remains in j, she will be counted as a knowledgeable worker for every year from 1994 to 1997.

which specifies that ζ_{jslt} contains measurement error ν_{jslt} and a productivity component ω^*_{jslt} (TFP) known to the firm but unobserved by the econometrician. The productivity component can be further divided into a firm-specific term, a term common to all firms in a given industry (μ_{st}) and a term common to all firms in a given locality (ϖ_{lt}). Equation (3) now becomes:

$$\ln(Y_{jslt}) = \beta_0 + \beta_L \ln(L_{jslt}) + \beta_K \ln(K_{jslt}) + \beta_M \ln(M_{jslt}) + \beta_H H_{jslt} + \mu_{st} + \varpi_{lt} + \omega_{jslt} + v_{jslt} \quad (5)$$

One major difficulty in estimating β_H in Equation (5) is that non HWFs may decide on their choice of H based on the realized firm-specific productivity shock (ω_{jslt}) unknown to the researcher[14]. In order to assess the relevance of this issue in my setting, in Section 5.1 I present estimates using control function methods in the spirit of the productivity literature (Olley and Pakes, 1996, henceforth OP; Levinsohn and Petrin, 2003, henceforth LP; Ackerberg, Caves, and Frazer, 2008, henceforth, ACF)

The number of knowledgeable workers may in principle also be correlated with productivity shocks happening in the future if workers can foresee them and apply for jobs at firms with better growth prospects. If such firms prefer to hire workers from good firms, these workers will have a higher probability of being chosen. To the extent that preferring workers from good firms can be explained through knowledge transfer from these firms, a positive correlation between H and the receiving firm's productivity shocks in $t+1$ does suggest a role for labor mobility as a channel for knowledge transfer, even though it will overestimate its importance (Stoyanov and Zubanov, 2012). To explore this possibility I adopt an IV strategy that I now describe.

Unobservable Future Shocks: Using the number of downsizing firms as IV In Section 5.2, I present estimates where I instrument for the number of knowledgeable workers in a non HWF with the number of *local* good firms in the *same 2-digit industry* that downsized in the previous period. Following a downsizing event at a HWF, it is more likely that a knowledgeable worker applies for job at local non HWFs because s/he is unemployed and does not want to relocate far away, and less likely that s/he does so because some particular non HWFs offer better prospects than the HWF at which the worker is employed. Put differently, in the scenario captured by the IV approach, the strategic mobility explanation is less likely to play a major role.

One can think of two main reasons why good firms may downsize in a particular year. First, good firms may get a bad draw from the distribution of product-market conditions.

[14]See Section A.I for a discussion using the first order condition with respect to H obtained from the optimization problem

Even though an inherent productivity advantage partly insulates the good firms from output shocks, sufficiently large shocks will pierce this insulation and induce the good firm to layoff workers. Alternatively, good firms may downsize in a particular year due to offshoring.

The basic intuition behind the IV approach is to consider moves from workers whose former employer downsized due to demand shocks or offshoring. While the timing of these moves is arguably exogenous, these workers may still decide which new employer to join among the set of non-HWFs. However, in small labor markets and specialized industries, workers may have a limited set of alternatives.

The choice of the instrument is based on the notion that geographic proximity plays an important role in determining worker mobility. In January 2012, I visited several Veneto firms and interviewed employees about the history of their enterprises and their current operations. I also conducted phone interviews with officials of employers' associations and chambers of commerce. My anecdotal evidence supports the idea that distance acts as a barrier for job mobility.[15] Moreover, in Section 4.2 I show descriptive evidence regarding the propensity of workers to change jobs within the same local labor market.

In the presence of product demand shocks or offshoring, using the number of downsizing firms as an instrument is invalid if it cannot be excluded from the causal model of interest (Equation 3). The identifying assumption of my IV strategy is therefore that the number of downsizing good firms is correlated with the causal variable of interest, H, but uncorrelated with any other unobserved determinants of productivity. In section 5.2 I discuss possible violations of the exclusion restrictions and describe my attempts at addressing such violations.

It is important to note that my instrument is an external 'z'-variable used in a production function framework. As pointed out in the survey Eberhardt and Helmers (2010), to date only past values of the regressors themselves or input prices have been used for instrumentation in the production function literature.

Alternative approach: proxy for future shocks To explore the possibility of future productivity shocks further I adapt the OP and LP approaches and also include in both t and $t+1$ (a) polynomial functions of capital and investment, and (b) polynomial functions of capital and materials, in an effort to proxy for shocks that may be anticipated by the workers. This is in the spirit of Stoyanov and Zubanov (2012) and assumes that hiring firms are also able to anticipate their productivity shocks and adjust their inputs accordingly. In Section 5 I provide the estimates and a longer discussion of such approach.

[15] In a phone interview, Federico Callegari of the Treviso Chamber of Commerce, reasoned out the role of geographic proximity: "I think distance matters a lot for workers' job mobility. When losing their job, workers tend to look for another job with a commuting time of maximum 20-30 minutes. Why? Because they want to go home during the lunch break!"

3.2.2 Unobserved Worker Quality

Another potential threat to identification is the fact that I do not observe labor quality. In order to investigate this issue, I obtain a proxy for worker ability θ_i, which I use in Section 5.3 to characterize knowledgeable workers, in the spirit of Balsvik (2011). I obtain the individual θ_i by procuring estimates of worker effects from wage equations where both firm and worker effects can be identified (Section 3.2.3 describes this estimation in detail).

3.2.3 Estimation of the Wage Model and Identification of Good Firms

Empirically, I identify potential high-productivity firms as high-wage firms (HWFs), i.e. those that pay a relatively high firm-specific wage premium. The use of alternative groupings of firms (good firms and non-good firms) based on output controlling for inputs and value added yields very similar results (5.3). The definition of potential high-productivity firms as high-wage firms (HWFs) employed in my baseline analysis is consistent with many recent models of frictional labor markets (e.g., Christensen et al., 2005), in which higher-productivity firms pay higher wages for equivalent workers. As I shall show below using balance sheet data, HWFs have significantly higher total factor productivity (TFP) and value added than other firms in my sample. There are three reasons why for the baseline results I define the good firms as HWFs and detect them using Social Security data rather than define the good firms directly as the high TFP or high value added ones and detect them using balance sheet data. First, the availability of worker-level Social Security data allows the introduction of measured individual characteristics and worker effects, something impossible to capture with firm level data from balance sheets. The estimated worker effects will also be helpful later in order to characterize knowledgeable workers and investigate the issue of unobserved labor quality when evaluating the productivity effect of labor mobility (recall the discussion in Subsection 3.2.2). Second, Social Security data are available for a longer period of time than the balance sheets, and therefore increase the precision of the categorization of firms into good and non-good groups. Third, since Social Security records are administrative data, measurement error is lower than in balance sheets.

Following Abowd, Kramarz and Margolis (1999, henceforth AKM), I specify a loglinear statistical model of wages as follows:

$$w_{ijt} = X_{it}'\beta + \theta_i + \psi_j + \varepsilon_{ijt} \tag{6}$$

where the dependent variable, the log of the average daily wage earned by worker i in firm j in year t, is expressed as a function of individual heterogeneity, firm heterogeneity, and measured time-varying characteristics.[16] The assumptions for the statistical residual ε_{ijt} are (a) $E[\varepsilon_{ijt}|i,j,t,x] = 0$, (b) $Var[\varepsilon_{ijt}|i,j,t,x] < \infty$ and (c) orthogonality to all other effects in the model. The presence of labor mobility in matched worker-firm data sets enables the identification of worker and firm effects.

A concern for estimation arises from the possibility of mobility based on the value of worker-firm match. In equation (6) ψ_j represents the wage premium paid to all employees in firm j, regardless of the features of the particular employees. Nevertheless, if the AKM exogenous mobility assumption is violated due to sorting based on the value of a worker-firm match component, and workers switch jobs to join firms to which they are better matched, then the wage premium would include a match component that would be specific to each employee-firm j pair, and no longer common across all employees in firm j. To test for such sorting, I perform three analyses: first, I look at wage changes for job siwtchers, second, I compare the AKM regression with a regression including match fixed effects, and third I examine the residuals from AKM. I present both analyses below. To summarize, I find little support for mobility based on the value of worker-firm match, consistent with Card, Heining and Kline (2013) and Macis and Schivardi (2013) results on German and Italian data respectively.[17] More details can be found in Section A.III.

For the baseline analysis, I identify good firms as those whose estimated firm fixed effects fall within the top third of all estimated firm effects.[18] Reults are very similar if I identify good firms as those whose estimated firm fixed effects fall within the top third of the estimated firm effects *within industry*. In Section 5.4 I remove the top third threshold and employ an alternative, continuous, measure of the receiving firm's exposure to knowledge, which exploits the differences between sending and receiving firms' quality, thus extending the analysis which works with a dummy indicating experience at a top third firm.

[16]The vector X'_{it} includes tenure, tenure squared, age, age squared, a dummy variable for manager and white collar status, and interaction terms between gender and other individual characteristics.

[17]More specifically, the absence of a mobility premium for the movers who remain in the same firm-effect quartile (Figure A.1) suggests that idiosyncratic worker-firm match effects are not the main driver of job mobility. Also, the symmetry between wage increases for job changers from low to high quartiles and the wage decreases for job changers in the opposite direction (Figure A.2), and the absence of notable systematic patterns in the distribution of residuals for particular types of matches (Figure A.3) are in line with the AKM model. Finally, the improvement in fit of a match fixed effects regression compared to the AKM model is very small. Notice that small match effects in wages do not automatically indicate small match effects in productivity however, as employees may have low bargaining power vis-à-vis their employers. In Section 5.3 I explore the possibility of match effects in productivity.

[18]Section 4.1 reports descriptive results as well as more details on the estimation procedure.

4 Data and Descriptive Statistics

The data used in this paper covers the region of Veneto, an administrative region in the Northeast of Italy with a population of around 5 million people (8 percent of the country's total). During the period of analysis (1992-2001), the labor market in Veneto has been characterized by nearly full employment, a positive rate of job creation in manufacturing and positive migration flows (Tattara and Valentini, 2010). The dynamic regional economy features a large presence of flexible firms, frequently organized in districts with a level of industrial value added greatly exceeding the national average. Within the district, larger lead firms often play an important coordinating role.[19] Manufacturing firms in Veneto specialize in metal-engineering, goldsmithing, plastics, furniture, garments, textiles, leather and shoes.[20] The manufacture of food and beverage, and wine and baked goods in particular, is also a prominent subsector.

My data set pools three sources of information: individual earnings records, firm balance sheets, and information on local labor markets (LLMs).[21] The earnings records come from the Veneto Workers History (VWH) dataset. The VWH has data on all private sector personnel in the Veneto region over the period 1975-2001. Specifically, it contains register-based information for virtually any job lasting at least one day. A complete employment history has been reconstructed for each worker.

Balance sheets starting from 1995 were obtained from AIDA (Analisi Informatizzata delle Aziende), a database circulated by Bureau Van Dijk containing official records of all incorporated nonfinancial Italian firms with annual revenues of at least 500,000 Euros. AIDA's balance sheets include firms' location, revenues, total wage bill, the book value of capital (broken into subgroups), value added, number of employees, value of materials and industry code. I use firm identifiers to match job-year observations for workers aged 16-64 in the VWH with firm financial data in AIDA for the period 1995-2001. Further details on the match and data restrictions I make, as well as descriptive information are provided in Section A.II.

Information on LLMs is 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. In 1991 the 518 municipalities or *comuni* in Veneto are divided into 51 LLMs.

[19] See Whitford (2001) for a discussion. The most famous industrial concentration is the eyewear district in the province of Belluno, where Luxottica, the world's largest manufacturer of eyeglasses, has production plants.

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

[21] The first two sources, combined for the period 1995-2001, have been used in the study on rent-sharing, hold-up and wages by Card, Devicienti and Maida (2014).

4.1 AKM Estimation and Descriptive Results

The method in Abowd, Creecy and Kramarz (2002) identifies separate groups of workers and firms that are connected via labor mobility in matched employer-employee data. When a group of workers and firms is connected, the group contains all persons who ever worked for any firm within the group and all firms at which any of the persons were ever employed. I run the grouping algorithm separately using VWH data from 1992 to 2000 for firms that could be matched in AIDA.[22] I then use the created group variable to choose the largest group as a sample for my fixed-effects estimation - Equation (6). Details on sample restrictions and descriptive information are provided in Appendix A.II.

I identify HWFs as those firms whose firm effects rank in the top third of the sample.Column 1 of Table 1 shows that HWFs pay on average 13 percent higher wages than non-HWFs[23]. For labor mobility to generate productivity benefits, a firm-specific advantage should be observed at good firms that could be the basis for knowledge transfer to other local firms.

I then estimate equations such as:

$$\ln O_{jslt} = \beta_0 + \beta_1 HWF_j + \mu_{st} + \varpi_{lt} + controls_{jt} + e_{jslt} \tag{8}$$

where the dummy HWF takes the value of 1 if firm j is classified as high-wage during the period 1992-2000 (the years over which the AKM estimates are obtained) and O_{jst} represents different firm-level outcomes over the period 1995-2001 (the years over which balance sheet data are available). The different firm-level outcomes are total factor productivity (output as dependent variable controlling for capital, material and labor inputs), value added, capital intensity (fixed assets as dependent variable controlling for firm size) and intangible capital intensity (intangible fixed assets - intellectual property, accumulated research and development investments and goodwill - as dependent variable controlling for firm size). Column 2-5 of Table 1 shows the results.

[TABLE 1 AROUND HERE]

In the Veneto manufacturing sector clear differences between HWFs and non HWFs emerge: HWFs feature on average 8 percent higher total factor productivity, 11 percent higher value added, 10 percent higher capital intensity and 27 percent higher intangible

[22] I experimented with other choices for the period of the AKM estimation, such as 1991-2000 or 1992-1999. Results are very similar.

[23] This finding emerges from the estimation of

$$w_{ijt} = X'_{it}\beta + \theta_i + \beta_1 HWF_j + \varepsilon_{ijt} \tag{7}$$

where the dummy HWF takes the value of 1 if firm j is classified as high-wage.

capital intensity. This evidence is important for establishing the potential for knowledge transfer in the region.

4.2 The Extent of Labor Mobility

For labor mobility to be a mechanism for transfer of knowledge, we must observe some workers moving from HWFs to other firms. This section documents the extent of labor mobility between HWFs and non-HWFs from 1992 to 2001. For this period, in my data set I observe around 52 thousands incidents of job change. These moves are categorized according to the direction of the flows in Table A.3. Column 1 shows that around 7,700 of these moves are from HWFs to non-HWFs. Column 2 of Table A.13 differentiates between moves within the same LLM and moves between LLMs. It shows that moves within the same LLM are more likely.[24] Column 3 of Table A.13 differentiates between moves within the same two-digit industry and moves between industries: around 35% of the moves from HWFs to non-HWFs are transitions to a non-HWF in the same industry in which the worker has HWF exprience. The remaining moves are transitions to a non-HWF in one of the nineteen 2-digit industries other than the one in which the worker has HWF experience.

Table A.4 shows the share of workers in non-HWFs with recent HWF experience. In 1995, only 0.5% of the employees in non-HWFs had recent HWF experience. In 2001 this share doubled to 1%. In terms of the potential for knowledge transfer, the relevant question is how these workers spread across the sample of non-HWFs. As shown in Table A.5, the share of firms employing workers with recent HWF experience is much greater than the share of such workers: around 18% in 1995 and around 29% in 2001. Therefore, during my sample period, an increasing percentage of firms employed workers with recent HWF experience.

4.2.1 Characteristics of Knowledgable Workers

Overall I observe 6539 unique knowledgable workers. Table A.6 shows their occupation. As regards to individual characteristics of the movers in my sample, Table A.7 shows that knowledgeable workers observed at non-HWF tend to be more likely to be male, white collars and managers than non-HWFs workers without recent experience at good firms. They also tend to be older.[25] For a comparison of the distribution of the estimated θ_i, see Section 5.3.

[24] In Section A.IV I further discuss the relation between geography and labor mobility.

[25] In terms of months of HWF experience, the minimum is 11 months, and the mean is 32 months.

5 Evidence on Worker Flows and Productivity

In this section I evaluate the extent to which non HWFs benefit from hiring workers from HWFs during the period 1995-2001. Details on sample restrictions and descriptive statistics for the variables used in the regression analysis are provided in Appendix A.II.

5.1 Estimates using OLS and control function methods

Table 2 shows the estimation results using OLS and control function methods. I cluster standard errors at the firm level. Coefficients associated with the H measure in Table 2 represent semielasticities because my variable of interest is not in logarithms. This choice for the baseline specification, which directly follows from Equation (2), is founded on the fact that H takes on the value 0 for a large number of observations (Figure 1). Thus, any possible transformation of the H measure could possibly affect the associated estimated parameters.[26]

[TABLE 2 AROUND HERE]

Column 1 reports estimate from the baseline OLS specification: the coefficient on H_{jst} is positive (0.03) and significant. In Column 2 and 3 I use control function methods in the spirit of the productivity literature (OP, LP) in order to address potential endogeneity arising from unobservable productivity shocks. [27] Although the point estimates of the coefficients for H_{jst} in the OP and LP specification are smaller than the baseline estimate, none of the specifications is qualitatively inconsistent with the empirical finding that non HWFs benefit from hiring workers from HWFs.

The extent to which non HWFs benefit from hiring workers from HWFs may be overestimated in Column 1-3 in the presence of productivity shocks happening in the future if workers can foresee them and apply for jobs in firms with better growth prospects (recall discussion in Section 3.2.1). In Section 5.2 I show results from the IV strategy. Columns 4 and 5 of Table 2 show the estimate from the alternative approach that I employ to address the issue of future productivity shocks: I add polynomial functions of capital and investments or capital and materials in t and $t+1$. These estimates also suggest that non HWFs benefit from knowledgeable workers by experiencing increased productivity.[28]

[26] Results using different functional forms are discussed in Section 5.3.

[27] Recall the discussion in Section 3.2.1. See Section A.V for a discussion of the OP and LP approaches, and the estimation details.

[28] That said, many components in the polynomial approximations are statistically significant, implying that these extra terms contribute in explaining the variation in productivity among firms. Notice the drop in observations due to the fact that I am using the leads of inputs (polynomials in $t+1$).

Overall, the main empirical result in this Subsection is that non HWFs benefit from hiring workers from HWFs. The point estimates suggest that the average effect of recruiting a knowledgeable worker on a non HWF's productivity is an increase of between 1.8 and 3 percent. This seems like a large effect. However, recall also that non-HWFs are quite small: the median number of employees at non HWFs is 33. Further, as many as 78 percent of non HWFs in a given year do not employ any worker of this type. Hiring one knowledgeable worker therefore implies a large change for most firms in our data. It is also instructive to evaluate the average magnitude of TFP change in monetary terms. This number can be calculated by multiplying the estimated percentage change by the mean value of non-HWF output. This calculation indicates that the increase in TFP due to hiring a worker from HWFs is associated with an increase in total output of 154-256 thousands of 2000 euros. As a further illustration of Table 2's estimates, the gain to a non-HWF hiring at the mean H (compared to an observationally identical firm that hired no-one) is equivalent to moving $3-5$ centiles up the productivity distribution for the median firm.[29]

5.2 IV Estimates

In this section I instrument for the number of knowledgeable workers using the lagged number of good local firms in the same 2-digit industry that downsized in the previous period. This exercise is motivated by the possibility of strategic mobility that I discussed in Section 3.2.1. The exclusion restriction is violated and $\widehat{\beta_H}^{IV}$ is biased upward if there are localized unobservable industry shocks that lead good firms to downsize and *positively* affect productivity at non HWFs. Below I discuss possible violations of the exclusion restrictions and describe my attempts at addressing such violations.

Turning to the details of the instrument, a downsizing firm must see an employment reduction larger than 1 percent *compared to the previous year's level.* The division of good firms into downsizing and non-downsizing firms according to this criterion is less sensible for small firms. Accordingly, I impose the additional condition that the decrease in the labor force is greater than or equal to three individuals.[30]

Table 3 shows the results from the IV estimation of Equation (3). Standard errors are clustered at the level of the LLM.

[TABLE 3 HERE]

[29] Stoyanov and Zubanov (2012) find that the productivity gains associated with hiring from more productive firms are equivalent to 0.35 percent per year for an average firm. Parrotta and Pozzoli (2012) find the impact of the recruitment of knowledge carriers on a firm's value added is an increase of 1%–2%. Balsvik (2011) finds that workers with MNE experience contribute 20% more to the productivity of their plant than workers without such experience.

[30] The instrumental variable is summarized in Table A.9, together with other variables constructed at LLM level that are used in the analysis.

The estimated coefficient of H in Column 1 is quite large (0.143). Recall the OLS estimates: the coefficient on knowledgeable workers is 0.03. A tentative explanation for the magnitude of the IV results is that the effect of knowledgeable workers may be heterogeneous across firms. If there are indeed heterogeneous effects of H on productivity, then consistent OLS measures the average effect of H on productivity across all firms, while Two Stage Least Squares (TSLS) estimates the average effect in the subset of firms that are marginal in the recruitment decision, in the sense that they recruit knowledgeable workers if and only if there exists excess local supply.[31] If the effect of knowledgeable workers on productivity is larger for non HWFs that are marginal in the recruitment decision, the TSLS estimates will exceed those of consistent OLS.

In principle, the IV estimates are also consistent with the idea that, since the good firms pay a relatively high firm-specific wage premium, workers who separate from a good firm may be of lower quality. I refer to this potential adverse selection problem as "lemons bias" (Gibbons and Katz, 1991). Lemons bias will tend to work against the finding of a positive effect of knowledgeable workers: in such scenario the OLS coefficient will be biased downward because of this negative selection. In practice, however, the IV standard errors are quite large (0.067) and prevent me from drawing definitive conclusions.

It is also important to emphasize that HWF downsize is not a likely event. For this reason the back-of-the-envelope calculations below (which take into account the probability of downsize) deliver a similar conclusion when using OLS and IV coefficient for $\widehat{\beta_H}$ to study the extent to which worker flows explain the productivity gains experienced by other firms when new highly productive firms are added to a LLM.

A concern for the validity of the exclusion restriction arises from the observation that the dependent variable in my econometric model is the *value* of output.[32] Unobserved shifts in local demand from HWFs to non-HWFs might simultaneously lead to (a) downsizing by HWFs, (b) higher output prices for non-HWFs, and (c) hiring of HWF employees by non-HWFs. The LLM-year effects control for local demand shocks, but localized unobservable *industry* shocks may still play a role. Consequently, in principle, it is possible that $\widehat{\beta_H}^{IV} > 0$ reflects higher output prices, rather than higher productivity due to labor mobility. I do not expect this to be a major factor in my context: manufacturing firms in my sample generally produce goods traded outside the LLM.[33] To further explore this possibility, in Column 2 I

[31] See Imbens and Angrist (1994) for a discussion. For a recent example, see Eisensee and Strömberg (2007).

[32] The theoretically correct dependent variable in a productivity study is the *quantity* of output, but, due to data limitations, this study (and virtually all the empirical literature on productivity) uses price multiplied by quantity.

[33] Imagine the extreme case of a non-HWF that produces a nationally traded good in a perfectly competitive industry. Its output prices would not increase disproportionately if the LLM experienced an increased demand for its good.

add a dummy taking value one if the industry produces goods that are not widely traded outside the LLM.[34]

Even when the level of tradability is controlled for, product demand effects might still be relevant and $\widehat{\beta_H}^{IV}$ might therefore be biased if an industry is strongly localized. In such a scenario the negative shock to the local HWF may lead to increased demand for the non HWF firm j even though the HWF and the non-HWF produce a tradable good. This is because, since most of the firms producing that particular good in Italy are in the same Veneto LLM, the non-HWF may experience an increase in demand, and hence in price, after the negative shock to a local HWF that is a direct competitor on the national market. To address this concern, I construct an index of industry localization as follows $r_s = (Italian\ Firms\ in\ s)/(Veneto\ Firms\ in\ s)$. Industries with low r have a relatively small number of firms outside the Veneto area. In Column 3 I enter r_s as additional regressor. The results in Column 2 and 3 are very similar to those in Column 1.

Finally, in column 4 I use an alternative definition of downsizing firms: a downsizing firm must see an employment reduction larger than 3 percent *compared to the previous year's level*.[35] The results are largely unchanged. Overall, the estimates show that the productivity effect of labor mobility is at least in part independent of unobserved future productivity shocks that are correlated with the propensity to hire workers with experience at highly productive firms.

5.3 Validity and Robustness

The main empirical result so far in the first part of the paper is that non HWFs benefit from hiring workers from HWFs. I now investigate the robustness of the estimates to various specifications and explore several possible alternative explanations for the estimated effects. Specifically, I (a) evaluate the role of unobserved worker quality, (b) present additional specifications addressing endogeneity concerns (time invariant firm level heterogeneity and time-varying unobservables), (c) discuss estimates using alternative groupings of firms, and (d) perform further robustness checks (value added specification, and investigation of the role of functional form assumptions).

Unobserved Worker Quality As explained above, potential threat to identification is the fact that I do not observe labor quality. To investigate this issue, Figures 2 shows a plot of the quantiles of the distribution of $\widehat{\theta_i}$ for the stayers at non-HWFs against the quantiles

[34] See Section A.VI for details.

[35] I keep the additional condition that the decrease in the labor force is greater than or equal to three individuals.

of the distribution of $\widehat{\theta}_i$ for the switchers from good firms. 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.[36] Since many points are on the left-hand side of the main diagonal, one tends to conclude that workers coming to non-HWFs are not positively selected on unobserved ability.

In Table A.10 I also show evidence, from regression analysis exploiting moves from non-HWFs, that it is not consistent with an explanation to my findings of a productivity effect of labor mobility based on unobservable worker quality.

Additional specifications addressing endogeneity concerns I start by addressing the issue of unobservables related with new hires. If workers who recently changed firms are more productive than stayers, the effect of newly hired workers with HWF experience may equally apply to newly hired employees without HWF experience. Also, the estimated productivity gains may be driven by better worker-firm matching rather than knowledge transfer. In order to explore these possibilities I first define medium-wage-firms (MWFs) as those whose estimated firm fixed effects from the AKM model fall between the 33th percentile and the 67th percentile of all estimated firm effects, and low-wage-firms (LWFs) as those whose estimated firm fixed effects fall below the bottom third. I then construct a new variable, denoted with N: the number of workers without recent experience at HWFs currently observed at MWF m. I then estimate for the sample of MWFs:

$$\begin{aligned}\ln(Y_{mslt}) &= \beta_0 + \beta_L \ln(L_{mslt}) + \beta_K \ln(K_{mslt}) + \beta_M \ln(M_{mslt}) + \\ &\quad + \beta_H H_{mslt} + \beta_N N_{mslt} + \mu_{st} + \varpi_{lt} + v_{mslt}\end{aligned}$$

In this specification, the identification of knowledge transfer relies on the differential effect of hiring an employee with recent HWF experience over hiring an employee from a LWF. By including both H and N, any potential bias caused by the correlation between unobservables and new hires is removed. Column 1 of Table A.10 shows the results. The coefficient of H is positive and significant. The coefficient of N is positive but much smaller. The difference in productivity premiums associated with the two types of newly hired workers is significant at the 1% level. This exercise, in the spirit of Balsvik (2011), can also be seen as a placebo test at firm-level and it suggests that the productivity effect attributed to knowledgeable workers is not associated with recently hired workers in general: large productivity gains linked to hiring seem to be realized only when new hires come from more productive firms. While I cannot completely rule out the possibility that at least some of the estimated effect reflects

[36] Both axes are in units of the estimated θ_i from equation 6 (vertical axis for stayers and horizontal axis for the hires from good firms). For a given point on the q-q plot, the quantile level is the same for both points.

better worker-firm matching, or switchers being more productive than stayers in general (i.e. regardless of the previous employment history), this evidence lends credibility to the knowledge transfer hypothesis. This evidence is also inconsistent with an explanation to our findings based on unobservable worker quality. If the estimates of a positive and significant β_H were driven by higher ability of workers moving from HWFs to MWFs, then β_N would be negative and significant, since hiring labor from non-HWFs would cause a decline in firm m's average worker quality and therefore deteriorate its productivity (Stoyanov and Zubanov, 2012).

Next, in Column 2 of Table A.10 I show estimates using the within-transformation in order to explore the possibility that the estimated productivity gains are due to time-invariant unobservables. This would be the case for instance if the (long-run) stable hiring patterns are due to certain management practices (Stoyanov and Zubanov, 2012). The estimates in Column 2 should be interpreted cautiously because the within estimator is known from practical experience to perform poorly in the context of production functions (Eberhardt and Helmer, 2010). [37] The problem of using the within-transformation is the removal of considerable information from the data, since only variation over time is left to identify parameters. Setting this concern aside, the results show a positive and significant coefficient on H.

Furthermore, considering the differences in observable characteristics documented in Appendix 4.2 between movers from HWFs and other workers at non HWFs, in Column 3 I augment Equation (3) with the share of females, managers, blue-collar and white-collar workers, and differently aged workers at each firm. The estimate of β_H in Column 3 is in line with the results from Table 2.

In Column 4, I include polynomial functions of capital, materials and number of employees in both t and $t+1$. This specification is in the spirit of the ACF approach. The estimates in Column 4, together with the IV results, and the estimate in Column 4 and 5 of Table 2 confirm the impression that the productivity effect of labor mobility is at least in part independent of unobserved future productivity shocks that are correlated with the propensity to hire workers with experience at highly productive firms.

Alternative groupings of firms As an additional sensitivity check, I identify potential good firms as high-TFP firms. Specifically, I estimate firm effects from a total factor productivity specification (i.e. one in which the dependent variable is output, and I control for inputs). I identify good firms as those whose estimated firm fixed effects fall within the top third of all estimated firm effects. The results, shown in Table A.11 are very similar to those

[37] Indeed, estimates in Column 2 indicate severely decreasing returns to scale, likely due to measurement error in the input variables, whose influence is exacerbated by the variable transformation.

in Table 2. I also experimented with a grouping of firm based on the estimated firm fixed effects in a value added specification. Results were largely unchanged.

Additional robustness checks Table A.12 shows results from further robustness checks. In Column 1, I use value added as an alternative measure of economic performance. Columns 2-5 investigate the role of functional form assumptions. Until now, I have presented results based on specifications where the intensity of potential knowledge transferred is measured by the number of H workers. In Column 2, I model this intensity as the share of workers with recent experience at good firms, dividing H by L.[38] In Column 3, I estimate:

$$\begin{aligned}\ln(Y_{jslt}) &= \beta_0 + \beta_L \ln(L_{jslt}) + \beta_K \ln(K_{jslt}) + \beta_M \ln(M_{jslt}) + \beta_{H_l} \log(H_{jslt}) + \\ +\delta 1(H &= 0)_{jslt} + \mu_{st} + \varpi_{lt} + v_{jslt}\end{aligned}$$

Compared to Equation (3), I replaced H_{jslt} with its logarithm, and I imposed $\log(H_{jslt}) = 0$ for the observations with $H_{jslt} = 0$. Plus, I added the dummy $1(H = 0)_{jslt}$ taking value 1 if the number of knowledgeable workers is equal to 0.

Column 4 allows the effect of each input to vary at the two-digit industry level. This specification accounts for the possibility that different industries use different technology or employ inputs of different quality. In Column 5, inputs are modeled with the translog functional form. My findings are robust to the different functional form assumptions in Columns 2 to 5. My main results are robust to each of these extensions.

5.4 Further Extensions

Results for Labor Mobility within and between Industry Sectors An interesting question is whether the knowledge embedded in workers is general enough to be applied in different industries: Column 1 of Table A.13 differentiates between workers with HWF experience moving within the same two-digit industry and workers moving between industries. The coefficient of both type of knowledgeable workers moving is significant and positive. This is consistent with knowledge transfer by labor mobility being able to overcome technology borders between industries.

[38] Since there may be measurement error in L, the number of employees in the AIDA data, a potential problem with such specification arises. Rewrite equation (3) as $\ln(\frac{Y_{jst}}{\theta L_{jst}}) = \beta_K \ln(K_{jst}) + \beta_M \ln(M_{jst}) + \beta_h h_{jst} + \mu_{st} + \varpi_{lt} + v_{jslt}$. Since $h = H/L$, a mechanical relantionship between h and the dependent variable may arise at time t. To address this issue, I use the share of H workers obtained from head counts in the Social Security dataset.

Results by Worker Occupation I now investigate whether new hires occupation influences the strength of the effect on the receiving firm productivity. I consider heterogeneity in knowledgeable workers' occupation both within their sending (HWF) and receiving (non-HWF). Specifically, knowledgeable workers are grouped into higher-skilled and lower-skilled occupations. The higher-skilled occupation category includes white collars and managers. The lower-skilled occupation includes blue collars and apprentices.

In Column 2 of Table A.13 the main variable of interest is disaggregated into two groups based on the occupation at the previous employer (HWF). In Column 3 it is disaggregated based on the occupation at the current employer (non-HWF). By and large, the estimates are consistent with the hypothesis that workers in higher-skilled occupation are better able to transfer knowledge. In both columns, coefficients on both variables are positive and significant, but the point estimate on the productivity effect is larger for switchers in higher-skilled occupations, with the differential impact being significant at conventional levels.

Continuous measure of the receiving firm's exposure to new knowledge In Column 4 of Table A.13 I employ an alternative, continuous, measure of the receiving firm's exposure to knowledge, which exploits the differences between sending and receiving firms, in the spirit of Stoyanov and Zubanov (2012). I thus extending the analysis above which has so far worked with a dummy indicating experience at a HWF. This new measure of firm's exposure to knowledge is calculated for each firm j hiring at time t as follows:

$$exposure_{j,t} = \sum_{i=1}^{G_{j,t}} D_{i,t}(\widehat{\psi}_i^s - \widehat{\psi}_j^r) \cdot G_{j,t}$$

where $\widehat{\psi}_i^s$ and $\widehat{\psi}_j^r$are the estimated AKM firm effects of the sending and receiving firms, $G_{j,t}$ is the number of new workers and $D_{i,t}$ is an indicator variable equal to one if $(\widehat{\psi}_i^s - \widehat{\psi}_j^r) > 0$ and zero otherwise. In words, the new measure is the difference in quality between the sending and receiving firm defined for each new worker i hired from more productive firms than the receiving firm j, multiplied by the number of such workers in j. The larger the value, the higher the exposure of the receiving firm to the knowledge coming from the sending firms. The estimates confirm that non HWFs benefit from hiring workers from HWFs.[39]

Lagged number of knowledgeable workers In Column 5 of Table A.13, I lag the number of workers with HWF experience. The coefficient is again positive and significant.

[39] A non-HWF hiring at the mean *exposure* is shown to feature 0.13 percent higher productivity compared to an observationally identical firm that hired no one.

6 Worker flows and agglomeration advantages

In this Section I assess the extent to which worker flows can explain the productivity advantages of firms located near other highly productive firms. In order to do so, I predict the change in productivity of a local firm j that is due to labor mobility following an event analogous to that studied by GHM. More specifically, the event I consider is an increase in the number of good firms such that the change in local output is comparable to the output of the average large plant whose opening is considered by GHM.[40]

I focus on a change in the number of good local firms belonging to *firm j's same industry.* This is motivated by the findings in Henderson (2003), Cingano and Schivardi (2004), Moretti (2004) and GHM that local spillovers are increasing in economic proximity.[41] An overview of my procedure is as follows. Denote the number of knowledgeable workers *moving within industry* observed at firm j with H^{ind}. As a first step, I estimate the effect on H_j^{ind} of a change in the number of good local firms belonging to firm j's same industry. Recall that if a worker is hired at time $t-g$, and has experience at a HWF between $t-g$ and $t-3$, she contributes to H count from year $t-g$ until t. This implies that H^{ind} exhibits a certain degree of persistence and suggests estimation of a dynamic model for the number of workers observed at firm j who have HWF experience in the same industry. In the second step, I predict the change in H^{ind} that each of the non HWFs in a LLM would experience if an output increase similar to the one considered by GHM were to occur, and I multiply the predicted change in H^{ind} by $\widehat{\beta_H}^{ind}$, the estimated coefficient on H^{ind} in my productivity regression. This product yields the predicted change in productivity due to worker flows for a given Veneto firm if its local industry were to experience an increase in output analogous to that considered by GHM.

In the final step, I compare my estimate of the predicted contribution of worker flows to productivity changes with GHM's estimate of the overall productivity effect. This comparison allows me to have a sense of the extent to which worker flows can explain the productivity gains experienced by other firms when high-productivity firms in the same industry are added to a local labor market. As I shall show below, a similar conclusion is reached if I use estimate of the overall productivity effect obtained in the Veneto data, rather than those in GHM.

[40] The large plants in GHM generated bidding from local governments, almost certainly because there was a belief of important positive effects on the local economy. GHM observe that the mean increase in TFP after the opening is (a) increasing over time and (b) larger if incumbent plants have the same industrial classification as 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. I think of the plants considered by GHM as "good" plants, and in order to simulate their experiment, I consider a change in the number of Veneto good firms such that the change in local output is comparable.

[41] Measures of economic links in these studies include (a) dummy indicating belonging to the same industry, (b) input and output flows and (c) indicators of technological linkages.

I will now discuss the issues related to the implementation of the first step, i.e. the estimation of the dynamic effect on H_j^{ind} of a change in the number of good firms in the same locality and industry.

6.1 A dynamic model for the number of knowledgeable workers

Consider a model of the form

$$H_{jlst}^{ind} = aH_{jsl,t-1}^{ind} + bGood\_Firms_{ls(j)t} + e_{jlst} \tag{9}$$

$$e_{jlst} = m_j + v_{jlst}$$
$$E[m_j] = E[v_{jlst}] = E[m_j v_{jlst}] = 0 \tag{10}$$

where $Good\_Firms_{ls(j)t}$ is the number of local good firms in the same industry of firm j. Recall that the subscript *ind* represent workers moving within industry. The disturbance term e_{jlst} has two orthogonal components: the firm effect, m_j and the idiosyncratic shock, v_{jlst}. Using OLS to estimate Equation (9) is problematic because the correlation between $H_{jsl,t-1}^{ind}$ and the firm effect in the error term gives rise to "dynamic panel bias" (Nickell, 1981). Application of the Within Groups estimator would draw the firm effects out of the error term, but dynamic panel bias would remain (Bond, 2002). Therefore I employ the first-difference transform, proposed by Arellano and Bond (1991):

$$\Delta H_{jlst}^{ind} = a\Delta H_{jsl,t-1}^{ind} + b\Delta Good\_Firms_{ls(j)t} + \Delta v_{jlst} \tag{11}$$

The firm effects have now disappeared, but the lagged dependent variable is still potentially endogenous as the $H_{jsl,t-1}^{ind}$ in $\Delta H_{jsl,t-1}^{ind} = H_{jsl,t-1}^{ind} - H_{jsl,t-2}^{ind}$ is correlated with the $v_{jls,t-1}$ in $\Delta v_{jlst} = v_{jls,t} - v_{jls,t-1}$. However, appropriately lagged values of the levels of the regressors remain orthogonal to the error and are available for use as instruments. Blundell and Bond (1998) show that under appropriate assumptions about the initial conditions, we can use appropriately lagged values of the differences of the regressors as instruments for the equation in levels. In the GMM system estimator, which I employ below, the orthogonality conditions for the differenced equation (11) are augmented by the orthogonality conditions for the level equation (9).[42]

[42] In principle, another challenge in estimating (11) is that firms in a given industry do not select their location randomly. Firms maximize profits and decide to locate where their expectation of the present discounted value of future profits is greatest. This net present value differs across locations depending on many factors, including transportation infrastructure, subsidies, etc. These factors, whose value may be different for firms in different industries, are unobserved, and they may be correlated with ΔH_{jlst}^{ind}. It should be noted, however, that a positive shock in LLM j and industry s such that there is entry of HWFs (i.e. an increase in $\Delta Good\_Firms_{ls(j)t}$) makes it *less* likely that a non-HWFs is going to hire from a good firm

Table 4 gives the results of estimating Equation (11) for the period 1992-2001. I include time dummies in order to remove universal time-related shocks from the errors.[43]

[TABLE 4 AROUND HERE]

Column 1 uses the system GMM estimator. I restrict the instrument set to lags 3 and longer, as suggested by the result of the Arellano-Bond test for serial correlation.[44] The regression shows a positive and significant coefficient of the number of good local firms. This is in line with the descriptive evidence discussed above of an important role of geographic and economic proximity in determining worker mobility. The column also shows a positive and significant coefficient for the lagged dependent variable. The p-value of the Hansen test for overidentifying restrictions does not suggest misspecification.

In Column 2, I estimate the model with two-step system GMM and Windmeijer (2005)-corrected cluster-robust errors.[45] In Column 1-2, for all variables only the shortest allowable lagged is used as instrument. In Column 3 and 4, I estimate the same specification in Column 1 including lags up to 4 and 5, respectively. The estimates in Columns 2 to 4 are similar to those in Column 1.

6.2 Back-of-the-envelope calculations

Having estimated the dynamic effect on H_j^{ind} of a change in $Good\_Firms_{ls(j)t}$, I can predict the changes in H, and hence in productivity, that a given non-HWF in Veneto would experience after an output increase similar to the one considered by GHM. As it turns out, the large manufacturing plants whose openings are studied by GHM are much larger than the typical good firm in Veneto.[46] In order to observe a change in local output comparable to the typical output increase caused by the opening of one large plant in GHM, a Veneto locality

in the same industry. This is because the shock is *good news* for good firms, so in principle it should make it less likely for the labor force at the good firms to experience a decrease, and in turn, it should make it less likely for a non-HWF to hire from a good firm. The bias introduced by the fact that good firms do not choose their location randomly is therefore likely to be downwards, and thus working against the finding of a positive effect of $\Delta Good\_firms_{ls(j)t}$ on ΔH^{ind}. In any case, $\Delta Good\_Firms_{ls(j)t}$ is treated as endogenous in the estimation.

[43] Since these specifications do not require information collected from AIDA balance sheets, the sample period is not restricted to post-1995 observations.

[44] Arellano and Bond (1991) develop a test for autocorrelation in the idiosyncratic disturbance term v_{jlst}. It checks for serial correlation of order l in levels by looking for correlation of order $l+1$ in differences. I employ this test below.

[45] See Roodman (2009) for a detailed discussion of two-step GMM and Windmeijer-correction.

[46] This is due both to the fact that new entrants in GHM are significantly larger than the average new plant in the United States and the fact that the Veneto region is characterized by the presence of small and medium-sized businesses, whose size is smaller than the typical firm in United States. See Section A.VII for descriptive statistics.

must experience an increase of 56 HWFs. This is the change in my back-of-the-envelope calculations.

The predicted change in H that a typical non-HWF would experience after 5 years, the time horizon considered in GHM, is then $\widehat{\Delta H}^{ind,5\ years} = 56 \cdot (b + ab + a^2b + a^3b + a^4b + a^5b)$. This change in H can be obtained using the estimates for a and b from Table 4. In order to obtain the predicted change in productivity, I use $\widehat{\beta_H}^{ind}$ from the estimation of Equation (3). The results using the different approaches (baseline OLS, OP, LP, polynomial functions of capital and investments or capital and materials in t and $t+1$, and IV) are shown in Table A.14. The predicted change in productivity attributable to worker flows five years the local output increase is then equal to $\widehat{\Delta TFP}^{ind,5\ years} = \widehat{\Delta H}^{ind,5\ years} \cdot \widehat{\beta_H}^{ind}$. In the case of the IV the number of new entrants is multiplied by the probability of downsizing. Table 5 provides a summation of the back-of-the-envelope calculations. The predicted change in productivity attributable to worker flows five years following a large local output increase ranges from 1.3% to 2.2% depending on the specification. The final step is to compare the magnitude of $\widehat{\Delta TFP}^{ind,5\ years}$ with GHM's estimate of the overall productivity effect caused by a local output increase. The increase in productivity estimated by GHM five years after the opening for incumbent plants in the same two-digit industry equals 17 percent. Hence, my back-of-the-envelope calculations suggest that worker flows explain 8-13% percent of the agglomeration advantages estimated by GHM, with the mean of the point estimates being 10%. Overall, these results suggest that worker flows explain an economically relevant proportion of the productivity gains experienced by other firms when HWFs in the same industry are added to a LLM.

6.3 Labor-market based knowledge spillovers and worker-firm matching

Recall my previous discussion of the agglomeration literature. A consensus has emerged that agglomeration economies can at least partially explain why firms cluster next to each other. Disagreement remains, however, over the sources of these agglomeration effects. In the above, I emphasized the possibility that knowledge is embedded in workers and diffuses when workers move between firms. The strong localized aspect of knowledge spillovers discussed in the agglomeration literature may thus arise from the propensity of workers to change jobs within the same local labor market.

Another explanation that has been proposed within the literature for the agglomeration of economic activity is the possibility of advantages deriving from thick labor markets. The argument is that agglomeration allows a better match between employer needs and worker skills, which may result in higher productivity (Helsley and Strange, 1990). In order to explore the relevance of this mechanism in the Veneto manufacturing sector context, I estimate a production function for non-HWF firm j in industry s and LLM l augmented with both the number of good firms and the number of non-good firms in industry s and LLM l.

$$
\begin{aligned}
\ln(Y_{jslt}) &= \widetilde{\beta}_0 + \widetilde{\beta}_L \ln(L_{jslt}) + \widetilde{\beta}_K \ln(K_{jslt}) + \widetilde{\beta}_M \ln(M_{jslt}) + b_G Good\_Firms_{js(j)t} + \quad () \\
&\quad + b_N Non - Good\_Firms_{js(j)t} + \varrho_{jslt} \qquad (12)
\end{aligned}
$$

I employ different estimation both OLS and GMM methods. When using GMM methods, both the number of good firms and the number of non-good firms are treated as endogenous (I experiment with different lags of the instruments). The results are shown in Table A.15. The number of good firms is positively and statistically significantly related to an increase in the productivity of non-HWF j. The coefficient of the number of non-good firms is negative and significant, or insignificant. The difference in productivity effects associated with each type of firm is significant. This exercise can also be seen as a placebo test at LLM-level and it suggests that the local productivity effect attributed to good firms is not associated with an increase in the size of the labor market in general: large productivity gains linked to changes in the number of firms seem to be realized only when the firms are good. Although I am not able to entirely discard the chance that at least part of the estimated impact reflects better worker-firm matching arising from a thicker labor market, this finding supports the hypothesis of labor-market based knowledge spillovers in the Veneto manufacturing sector context.

Comparison with the results in GHM It is instructive to compare (a) the prediction from Table A.15 regarding the overall change in productivity after an increase in local output, with (b) the effect found by GHM. Recall that in order to observe a change in local output comparable to the typical output increase caused by the opening of one large plant in GHM, a Veneto locality must experience an increase of 56 HWFs. The predicted change of such increase in the number of good firms based on Table A.15 is in the range 18-33%, quite consistent with the estimate in GHM (17%).

7 Conclusions

Localized knowledge spillovers are a common explanation for the productivity advantages of agglomeration.[47] Nevertheless, as pointed out by Combes and Duranton (2006), if infor-

[47] The availability of specialized intermediate inputs, the sharing of a labor pool, and better matching have also garnered attention in the literature's attempt to explain agglomeration economies.

mation can easily flow out of firms, the question of why the effects of spillovers are localized must be clarified.

This paper directly examined the role of labor mobility as a mechanism for the transfer of efficiency-enhancing knowledge and evaluated the extent to which labor mobility can explain the productivity advantages of firms located near other highly productive firms. The underlying idea is that knowledge is embedded in workers and diffuses when workers move between firms. The strong localized aspect of knowledge spillovers discussed in the agglomeration literature may thus at least in part arise from the propensity of workers to change jobs within the same local labor market.

In order to empirically assess the importance of labor-market based knowledge spillovers, I used Social Security earnings records and detailed financial information for firms from the Veneto region of Italy. The main empirical task is to show that the observed associations between labor mobility and productivity are at least in part causal. I implement several strategies to support a causal explanation, which include control function methods, IV strategy, and placebo tests. While none of these strategies is completely conclusive in regard to identification, together they give evidence that is consistent with a casual interpretation of the observed labor mobility effects and inconsistent with the plausible alternative explanations.

The empirical evidence presented using this unique dataset points to the concrete possibility that agglomeration of economic activity creates important productivity advantages at the local level. The productivity benefits of a non-HWF from being located in a cluster with a large number of good firms rest with the opportunities to hire workers whose knowledge was gained in good firms. Such knowledge can be successfully adapted internally. More specifically, the regression analysis showed that hiring a worker with HWF experience increases the productivity of other (non-HWF) firms. Back-of-the-envelope calculations indicated that worker flows explain a significant portion of the productivity gains experienced by other firms when HWFs in the same industry are added to a local labor market.

References

Abowd, J.M., F. Kramarz, and D.N. Margolis. 1999. "High-wage Workers and High-wage Firms." *Econometrica*, 67(2): 251–333.

Abowd, J.M., R.H. Creecy, and F. Kramarz. 2002. "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data." Center for Economic Studies, U.S. Census Bureau Longitudinal Employer-Household Dynamics Technical Papers.

Alfaro, L., and M.X. Chen. 2014. "The Global Agglomeration of Multinational Firms." *Journal of International Economics, forthcoming.*

Arellano, M., and S. Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *The Review of Economic Studies*, 58(2): 277–297.

Balsvik, R. 2011. "Is Labor Mobility a Channel for Spillovers from Multinationals? Evidence from Norwegian Manufacturing." *The Review of Economics and Statistics*, 93(1): 285–297.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen. 2013. "Identifying Technology Spillovers and Product Market Rivalry." *Econometrica*, 81(4): 1347–1393.

Blundell, Richard, and Stephen Bond. 1998. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of econometrics*, 87(1): 115–143.

Bond, S.R. 2002. "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice." *Portuguese Economic Journal*, 1(2): 141–162.

Brusco, S. 1983. "The Emilian Model: Productive Decentralisation and Social Integration." *Cambridge Journal of Economics*, 6(2): 167–184.

Card, David, Francesco Devicienti, and Agata Maida. 2014. "Rent-sharing, holdup, and wages: Evidence from matched panel data." *The Review of Economic Studies*, 81(1): 84–111.

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *The Quarterly Journal of Economics*, 128(3): 967–1015.

Christiansen, B.J., R. Lentz, G. Mortensen, G. Neumann, and A. Werweck. 2005. "Job Separations and the Distribution of Wages." *Journal of Labor Economics*, 23: 31–58.

Cingano, F., and F. Schivardi. 2004. "Identifying the Sources of Local Productivity Growth." *Journal of the European Economic Association*, 2(4): 720–744.

Combes, P. P., and G. Duranton. 2006. "Labour Pooling, labour Poaching, and Spatial Clustering." *Regional Science and Urban Economics*, 36(1): 1–28.

Combes, P.P., G. Duranton, L. Gobillon, D. Puga, and S. Roux. 2012. "The productivity advantages of large cities: Distinguishing agglomeration from firm selection." *Econometrica*.

Cooper, D. P. 2001. "Innovation and Reciprocal Externalities: Information Transmission Via Job Mobility." *Journal of Economic Behavior and Organization*, 45(4): 403–425.

Dasgupta, Kunal. 2012. "Learning and knowledge diffusion in a global economy." *Journal of International Economics*, 87(2): 323–336.

Duranton, G., and D. Puga. 2004. "Micro-foundations of Urban Agglomeration Economies." *Handbook of Regional and Urban Economics*, 4: 2063–2117.

Eberhardt, M., and C. Helmers. 2010. *Untested Assumptions and Data Slicing: A Critical Review of Firm-Level Production Function Estimators.* Department of Economics, University of Oxford.

Eisensee, T., and D. Strömberg. 2007. "News Droughts, News Floods, and U.S. Disaster Relief." *The Quarterly Journal of Economics*, 122(2): 693–728.

Ellison, G., E. L. Glaeser, and W. Kerr. 2010. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns." *American Economic Review*, 3: 1195–1213.

Fosfuri, A., M. Motta, and T. Rønde. 2001. "Foreign Direct Investment and Spillovers Through Workers' Mobility." *Journal of International Economics*, 53(1): 205–222.

Gibbons, R., and L.F. Katz. 1991. "Layoffs and Lemons." *Journal of Labor Economics*, 9(4): 351–380.

Glass, A.J., and K. Saggi. 2002. "Multinational Firms and Technology Transfer." *The Scandinavian Journal of Economics*, 104(4): 495–513.

Greenstone, M., R. Hornbeck, and E. Moretti. 2010. "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plants Openings." *Journal of Political Economy*, 118: 536–598.

Griliches, Zvi, and Jacques Mairesse. 1995. "Production functions: the search for identification." National Bureau of Economic Research.

Helsley, R.W., and W.C. Strange. 1990. "Matching and Agglomeration Economies in a System of Cities." *Regional Science and Urban Economics*, 20(2): 189–212.

Henderson, J.V. 2003. "Marshall's Scale Economies." *Journal of Urban Economics*, 53(1): 1–28.

Henry, Nick, and Steven Pinch. 2000. "Spatialising knowledge: placing the knowledge community of Motor Sport Valley." *Geoforum*, 31(2): 191–208.

Imbens, G.W., and J.D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, 62(4): 467–476.

Levinsohn, J., and A. Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of Economic Studies*, 70(2): 317–42.

Macis, Mario, and Fabiano Schivardi. 2013. "Exports and Wages: Rent Sharing, Workforce Composition or Returns to Skills?"

Markusen, J.R. 2001. "Contracts, Intellectual Property Rights, and Multinational Investment in Developing Countries." *Journal of International Economics*, 53(1): 189–204.

Moretti, E. 2004. "Estimating the External Return to Higher Education: Evidence from Cross-sectional and Longitudinal Data." *Journal of Econometrics*, 120: 175–212.

Moretti, E. 2011. "Local Labor Markets." *Handbook of Labor Economics*, 4: 1237–1313.

Nickell, S. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica*, 49: 1417–1426.

Olley, G.S., and A. Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64(6): 1263–97.

Parrotta, P., and D. Pozzoli. 2012. "The Effect of Learning by Hiring on Productivity." *The RAND Journal of Economics*, 43(1): 167–185.

Petrin, A., B.P. Poi, and J. Levinsohn. 2004. "Production Function Estimation in STATA Using Inputs to Control for Unobservables." *Stata Journal*, 4: 113–123.

Piore, MJ. 2009. "Conceptualizing the Dynamics of Industrial Districts." In *The Handbook of Industrial Districts*. Cheltenham, Edward Elgar.

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

Poole, Jennifer P. 2013. "Knowledge transfers from multinational to domestic firms: evidence from worker mobility." *Review of Economics and Statistics*, 95(2): 393–406.

Roodman, D. 2009. "How To Do Xtabond2: An Introduction to Difference and System GMM in Stata." *Stata Journal*, 9(1): 86–136.

Rosenthal, S.S., and W.C. Strange. 2003. "Geography, Industrial Organization, and Agglomeration." *Review of Economics and Statistics*, 85(2): 377–393.

Rosenthal, S.S., and W.C. Strange. 2004. "Evidence on the Nature and Sources of Agglomeration Economies." *Handbook of Regional and Urban Economics*, 4: 2119–2171.

Saxenian, A. 1994. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128.* Harvard University Press.

Stoyanov, A., and N. Zubanov. 2012. "Productivity Spillovers Across Firms through Worker Mobility." *American Economic Journal: Applied Economics*, 4(2): 168–198.

Tattara, G., and M. Valentini. 2010. "Turnover and Excess Worker Reallocation. The Veneto Labour Market between 1982 and 1996." *Labour*, 24(4): 474–500.

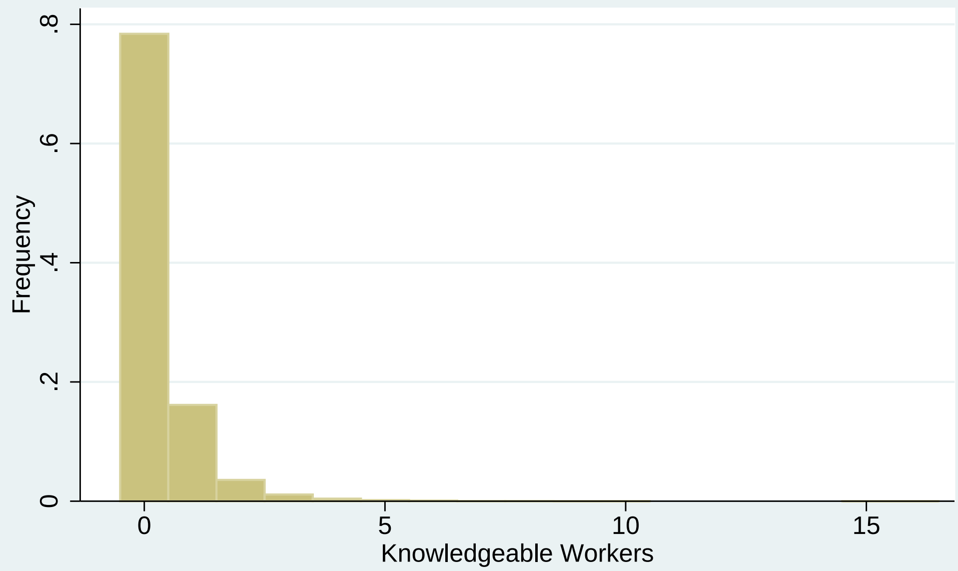
Trigilia, C. 1990. "Work and Politics in the Third Italy's Industrial Districts." *Industrial Districts and Inter-Firm Co-operation in Italy, Geneva: International Institute for Labor Studies*, F. Pyke, G. Becattini and W. Sengenberger (eds.): 160–184.

Whitford, J. 2001. "The Decline of a Model? Challenge and Response in the Italian Industrial Districts." *Economy and Society*, 30(1): 38–65.

Windmeijer, Frank. 2005. "A finite sample correction for the variance of linear efficient two-step GMM estimators." *Journal of econometrics*, 126(1): 25–51.

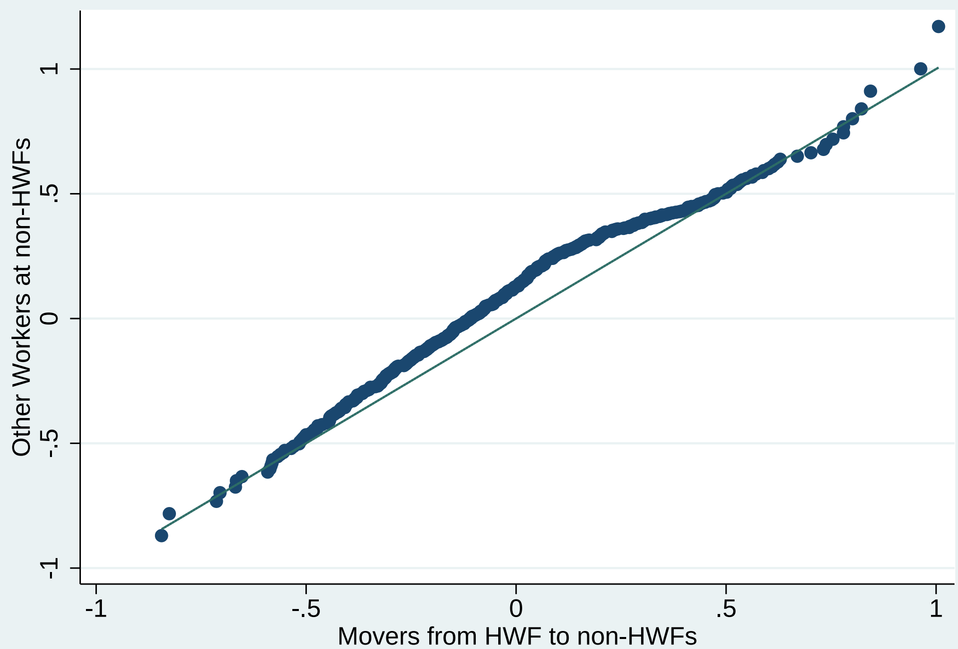
Yasar, M., R. Raciborski, and B. Poi. 2008. "Production Function Estimation in STATA Using the Olley and Pakes method." *Stata Journal*, 8(2): 221.

Figure 1: *Distribution of H Workers across Firms*



Number of Knowledgeable Workers observed at non-HWFs. Firm-year observations 1995-2001. I define a worker as having recent HWF-experience in year t, if he or she is observed working in a HWF for one or more of the years t-3 to t-1. If a worker is hired at time t-g, and has experience at a HWF between t-g and t-3, he or she contributes to the H count from year t-g until t.

Figure 2: *Q-Q Plot:Worker Effects*



Quantile-Quantile Plot of Worker Effects in1995. Plots for all other years are very similar

Table 1: Characteristics of HWFs

	(1) Individual Wage	(2) TFP	(3) Value Added	(4) K	(5) Intangible K
HWF	0.130	0.080	0.105	0.101	0.274
	(0.003)	(0.008)	(0.010)	(0.025)	(0.043)
Observations	1837597	26657	26587	26674	24450
Adj. R-squared	0.912	0.920	0.800	0.496	0.210

The Table shows that in the Veneto manufacturing sector clear differences between HWFs and non HWFs emerge. This evidence is important for establishing the potential for knowledge transfer in the region. The dummy HWF takes value 1 if the firm is classified as high-wage during the period 1992-2000 (the years over which the AKM estimates are obtained). Dependent Variables are in logs. In Column 1 the dependent variable is individual wage. In Column 2-5 the different firm-level outcomes are total factor productivity (output as dependent variable controlling for capital, material and labor inputs), value added, capital intensity (fixed assets as dependent variable controlling for firm size) and intangible capital intensity (intangible fixed assets - intellectual property, accumulated research and development investments and goodwill - as dependent variable, controlling for firm size). Output, Value Added and Capital variables are in 1000's of 2000 euros and are measured over the period 1995-2001 (the years over which balance sheet data are available). Standard errors (in parentheses) clustered by firm.

Table 2: Knowledgeable Workers and Productivity in non-HWFs, 1995-2001

	(1) OLS	(2) OP	(3) LP	(4) Inv-Cap Interactions t,t+1	(5) Mat-Cap Interactions t,t+1
log(capital)	0.092	0.087	0.148		
	(0.005)	(0.019)	(0.010)		
log(materials)	0.583	0.587		0.617	
	(0.007)	(0.007)		(0.012)	
log(employees)	0.223	0.225	0.202	0.187	0.177
	(0.006)	(0.010)	(0.006)	(0.014)	(0.006)
H workers	0.030	0.018	0.021	0.018	0.022
	(0.003)	(0.004)	(0.003)	(0.006)	(0.003)
Observations	17158	6635	17158	2963	13540
Adj. R-squared	0.931			0.940	0.952

Dependent variable: Log(Output). Standard errors (in parentheses) clustered by firm. H workers is the number of workers with HWF experience currently observed at non-HWFs. Column 1 reports estimates from the baseline specification. Column 2 implements the procedure in Olley and Pakes (1996). Column 3 implements the procedure in Levinsohn and Petrin (2003). Column 4 adds a third-degree polynomial function of log capital and log investment and the interaction of both functions in t and t+1. Column 5 includes the same controls as col. 4 but replaces log investment with log materials.

Table 3: Knowledgeable Workers and Productivity in non-HWFs, IV Estimates 1995-2001

	(1) Baseline	(2) tradability	(3) localization	(4) Larger drop in L
H workers	0.143	0.143	0.147	0.172
	(0.067)	(0.067)	(0.066)	(0.083)
log(capital)	0.085	0.085	0.084	0.083
	(0.008)	(0.008)	(0.008)	(0.008)
log(materials)	0.575	0.575	0.575	0.573
	(0.013)	(0.013)	(0.013)	(0.014)
log(employees)	0.204	0.204	0.203	0.199
	(0.010)	(0.010)	(0.010)	(0.011)
Observations	17566	17566	17566	17566
Adj. R-squared	0.903	0.902	0.902	0.906
Fstat, instrum., 1st stage	19.09	18.82	22.60	16.90

Dependent variable: Log(Output). Standard errors (in parentheses) clustered by LLM (47). Regressions include industry-year interaction dummies and LLM-year interaction dummies. Column 1 reports IV estimates using the lagged number of downsizing local good firms in the same 5-digit industry. A good firm is considered as downsizing if the drop in L is larger than 1 percent. The decrease in the labor force must also be greater than or equal to three individuals. Column 2 adds an indicator of the importance of local demand, namely a dummy taking value 1 if the 4-digit industry produces goods that are not widely traded outside the LLM. Column 3 controls for an index of industry localization, namely the ratio between the number of firms in Veneto and total Italian firms in a given 4-digit industry. In Column 4 a good firm is considered as downsizing if the drop in L is larger than 3 percent. The controls are the same as in Column 3.

Table 4: Number of local HWFs in same Industry and Knowledgeable Workers moving within industry, System GMM Estimates, 1992-2001

	(1) Baseline	(2) Twostep	(3) Lags up to 4	(4) Lags up to 5
lag(H from same Ind)	0.144	0.136	0.147	0.159
	(0.0719)	(0.0653)	(0.0717)	(0.0660)
Local HWFs in same Ind	0.009	0.009	0.010	0.009
	(0.0012)	(0.0012)	(0.0015)	(0.0014)
Observations	25688	25688	25688	25688
AR(1)z	-2.124	-2.436	-2.172	-2.337
AR(2)z	-6.062	-6.339	-6.057	-6.010
AR(3)z	0.304	0.196	0.337	0.460
HansPv	0.321	0.321	0.607	0.941

Dependent variable: 'H from same Ind', the number of H workers moving within Industry. Standard errors (in parentheses) clustered by LLM. Regressions include year dummies. The variable 'Local HWFs in same industry' is treated as endogenous. Column 1 reports the baseline System GMM results. Column 2 estimates the model with two-step System GMM with Windmeijer-corrected standard errors. I restrict the instrument set to lags 3 and longer. In Column 1-2 for all variables only the shortest allowable lagged is used as instrument. In Column 3 and 4 lags up to 4 and 5 are used, respectively. AR(1)z, AR(2)z and AR(3)z: Arelanno and Bond (1999) test of first, second and third order serial correlation, distributed as N(0,1). HansPv: p-value of Hansen test of overidentifying restrictions.

Table 5: Worker flows and agglomeration advantages

	(1) OLS	(2) OP	(3) LP	(4) Inv-Cap Interactions t,t+1	(5) Inv-Mat Interactions t,t+1	(6) IV
$\widehat{\beta_H}^{ind}$	0.036	0.037	0.031	0.022	0.026	0.121
Probability of HWF downsize	...	...	...	...	...	0.178
$\widehat{\Delta TFP}^{ind,\ 5\ years} = \widehat{\Delta H}^{ind,\ 5\ years} * \widehat{\beta_H}^{ind}$	0.021	0.022	0.018	0.013	0.015	0.013
$\widehat{\Delta TFP}^{ind,\ 5\ years}$/overall agglom. effect	0.125	0.128	0.107	0.076	0.090	0.075

This table provides a summation of the predicted change in productivity that is attributable to worker flows five years following a local output increase. The predicted changes are calculated for each of the different functional forms (i.e. $\widehat{\beta_H}^{ind}$ is obtained using OLS, OP, LP, polynomial functions of both capital and investments and capital and materials, and IV). Simulating results to correspond to the large plant opening results found in GHM such that one large plant opening is equivalent to 56 small Veneto plants, this table provides evidence that worker flows explain an important portion of the agglomeration advantages found in GHM.

Appendix

A.I The firm problem, and the role of productivity shocks

Consider a firm j 's optimization problem in each time period:

$$\underset{L_j,K_j,M_j}{max}\, \pi_j = A_j[{L_j}^{\alpha} K_j^{\gamma} M_j^{\lambda}]^{\delta} - w_0 N_j - w_1 H_j - \rho K_j - \tau M_j - f(H_j) \tag{13}$$

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); the term $f(H_j)$ captures the cost of recruiting knowledgeable workers (due to search frictions for instance). I assume $f' > 0$, $f" > 0, f'(0) = 0$.[48] For readability I drop the time subscript (since all terms are contemporaneous), and industry and locality subscripts. The corresponding first-order condition for H_j is[49]

$$\frac{\partial \pi}{\partial H_j} = \beta_{H_j} D e^{\beta_{H_j} H_j} [(N_j + H_j)^{\alpha} Q_j]^{\delta} + \delta D e^{\beta_{H_j} H_j} [(N_j + H_j)^{\alpha} Q_j]^{\delta - 1} \times \left(\alpha (N_j + H_j)^{\alpha - 1} Q_j\right) - w_1 - f'(H_j) = 0 \tag{14}$$

where $Q_j = K_j^{\gamma} M_j^{\lambda}$. Solving (14) for D, we have that

$$D = \frac{w_1 + f'(H_j^*)}{(K_j^{\gamma} M_j^{\lambda})^{\delta} e^{\beta_{H_j} H_j} \left[\beta_{H_j} (H_j^* + N_j)^{\alpha} + \delta\alpha (H_j^* + N_j)^{\alpha - 1}\right]}$$

Taking first-order conditions with rispect to N and M and combining them with the above expression yields[50] :

$$D = \frac{w_1 + f'(H)}{\left(\frac{w_0}{\tau}\right)^{\alpha - 1} K^{\gamma\delta} e^{\beta_H H} \left(w_0 \left(\frac{\tau}{w_0}\right)^{\lambda\delta} \frac{1}{AK^{\gamma\delta}}\right)^{\frac{\alpha}{\lambda\delta + \delta\alpha - 1}} \left[\beta_H \left(w_0 \left(\frac{\tau}{w_0}\right)^{\lambda\delta} \frac{1}{AK^{\gamma\delta}}\right)^{\frac{1}{\gamma\delta + \delta\alpha - 1}} \frac{w_0}{\tau} + \delta\alpha\right]} \tag{15}$$

If $f(\cdot)$ is convex enough, then this equation gives us the optimal H_j^*, with $\frac{dH_j^*}{dD} > 0$, i.e., there is a positive relationship between D and H_j.[51] In such case, when employing OLS to estimate Equation (5) without accounting for the existence of ω_{jslt}, the bias induced by endogeneity between H_j and ω_{jslt} is positive (positive productivity shocks translate into higher probability to hire from HWFs), implying that the coefficient estimate will be biased upward ($\widehat{\beta_{H_j}} > \beta_{H_j}$).[52]

[48] This in order to avoid corner solutions (i.e. the firm hiring only knowledgeable workers).

[49] Recall that $L_j = H_j + N_j$ and $A_j = D_j e^{\beta_H H_j}$

[50] I consider K as a state variable, in line with the recent productivity literature.

[51] Essentially, because H_j affects productivity, one needs to ensure that the marginal recruiting cost $f'(\cdot)$ increases sufficiently fast so that the firm has decreasing returns to scale.

[52] Recall that $\ln(D_j) = \beta_0 + \zeta_{jslt}$ and $\zeta_{jslt} = \omega_{jslt} + \mu_{st} + \varpi_{lt} + \nu_{jslt}$ where the firm-specific productivity shock (ω_{jslt}) is unknown to the researcher.

A.II Sample Restrictions, AKM Estimation Details and Additional Descriptive Information

Following Card, Devicienti, Maida (2014), I use firm identifiers to match job-year observations for workers aged 16-64 in the VWH with firm financial data in AIDA for the period 1995-2001. The match rate is fairly high: at least one observation in the VHW was found for over 95 percent of the employers in the AIDA sample, and around 50 percent of employees observed in the VWH between 1995 and 2001 can be matched to an AIDA firm. Most of the nonmatches seem to be workers of small firms that are omitted from AIDA. In sum, I was able to match at least one employee for around 18,000 firms, or around 10 percent of the entire universe of employers contained in the VWH.[53] From this set of potential matches I execute two exclusions to obtain my estimation sample for Equation (6). First, I remove all workers outside manufacturing. Next, I exclude job-year observations with remarkably high or low values for wages (I trim observations outside the 1 percent - 99 percent range).

The method in Abowd, Creecy and Kramarz (2002) identifies separate groups of workers and firms that are connected via labor mobility in the data. I run the grouping algorithm separately using VHW data from 1992 to 2000 for firms that could be matched in AIDA. I then use the created group variable to choose the largest group as the sample for my fixed-effects estimation. Figure A.4 shows the distribution of estimated firm effects.[54] I identify HWFs as those firms whose firm effects rank in the top third of the sample.[55] Table A.1 summarizes the sample of HWFs. Figure A.5 shows the geographical variation in the number of HWFs across LLMs. Table A.2 compares HWFs and non HWFs in terms of workforce characteristics. [56]Table A.8 summarizes the sample of non HWFs used in the main firm-level analysis – equation (3).[57] The main analysis is performed over the period for which balance sheet data are available (1995-2001).

[53] Average firm size for the matched jobs sample (36.0 workers) is considerably larger than that for total employers in the VWH (7.0 workers). Mean daily wages for the matched observations are also greater, while the fractions of under 30 and female employees are lower.

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

[55] Reults are very similar if I identify good firms as those whose estimated firm fixed effects fall within the top third of the estimated firm effects *within industry*.

[56] Notice that since the specifications in Table A.2 do not require information collected from AIDA balance sheets, the sample period is not restricted to post-1995 observations.

[57] In order to obtain this firm-level estimation sample I first remove the HWFs. From this non-HWF sample I remove (a) firms that close during the calendar year and (b) firm-year observations with remarkably high or low values (outside the 1% - 99% range) for several key firm-level variables, such as total value of production, number of employees, capital stock and value of materials, share of workers with recent experience at good firms (obtained dividing H by L) (c) firms in LLM with centroids outside Veneto (3 LLMs). I then attempt to reduce the influence of false matches, particularly for larger firms, by implementing a strategy of Card, Devicienti and Maida (2014) to eliminate 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 a firm's AIDA balance sheet and the number found in the VWH is larger than 100.

A.III Mobility based on the value of worker-firm match

To test for mobility based on the value of worker-firm match (see discussion in Section 3.2.3), I follow Card, Heining and Kline (2013) and Macis and Schivardi (2013) and perform three analyses: first, I look at wage changes for job siwtchers, second, I compare the AKM regression with a regression including match fixed effects, and third I examine the residuals from AKM. I present these analyses below.

A.III.1 Wage changes for job switchers

I consider all job switchers in the years 1992-2001 with at least two consecutive years at the old and new employer. I then categorise the source and destination jobs based on the quartiles of the estimated ψ_j's. I form sixteen cells based on quartiles of source and destination, and calculate average wages of switchers in each cell in the two years before the switch and the two years after the switch. Under the exogenous mobility assumption, workers who move between employers that pay comparable wages should not experience any wage change. Further, workers who move from a "high ψ" to a "low ψ" employer should experience a wage loss and workers who move in the reverse way a wage gain. Moreover, the wage decrease for the former set and the wage increase for the latter set be approximately symmetrical – the "ψ" lost by one set should be approximately the same of that lost by the other set. Figure A.2 and A.3 show patterns consistent with such implications of the exogenous mobility assumption. The absence of a mobility premium for the movers who remain in the same firm-effect quartile suggests that idiosyncratic worker-firm match effects are not the main driver of job mobility. Also, the symmetry between wage increases for job changers from low to high quartiles and the wage decreases for job changers in the opposite direction are in line with the AKM model

A.III.2 Contrast of AKM and match fixed effects regression

If match effects are significant, a model with worker-firm fixed effects should out-perform the AKM model as regards to statistical fit. I find that for the AKM: Adj R-squared = 0.91, Root MSE = 0.077; for the match fixed effects regression Adj R-sq = 0.92, Root MSE = 0.084. Even if these results show the presence of a match component in wages, the improvement in fit of a match fixed effects regression compared to the AKM model is very small

A.III.3 Analysis of the residuals from AKM

I also analyze the residuals from the AKM regression. Specifically, I form deciles based on the estimated worker effects and firm effects, and calculate average residuals in each of the

100 worker x firm decile cell, to examine whether there are any notable systematic patterns in the distribution of residuals for particular types of matches. The absence of such patterns (Figure A.3) supports the conclusion that in the Veneto manufacturing sector context, the additively separable firm and worker effects obtained from the AKM model represent sound measures of the unobservable worker and firm components of wages.

A.IV Geography and Labor Mobility: Further Discussion

There exist at least two reasons why geographic proximity might be important for observed worker flows. First, distance may act as a barrier for workers' job mobility because of commuting costs or idiosyncratic preferences for location. Descriptive statistics in Combes and Duranton (2006) show that labor flows in France are mostly local: about 75% of 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 kilometers. In Dal Bo', Finan and Rossi (2013), randomized job offers produce causal estimates of the effect of commuting distance on job acceptance rates. Distance appears to be a very strong (and negative) determinant of job acceptance: applicants are 33% less likely to accept a job offer if the municipality to which they are assigned is more than 80 kilometers away from their home municipality. The estimates in Manning and Petrongolo (2013) also suggest a relatively fast decay of job utility with distance. Another reason geographical proximity may be an important determinant of job mobility is that the firm's informational cost of identifying the "right" employee are larger across localities than within them. A similar argument can be made for the informational costs for workers.

A.V Discussion of OP and LP approaches, and estimation details

OP construct an explicit model for the firm's optimization problem in order to obtain their production function estimator. Essentially, the authors address the issue of endogeneity of inputs by inverting the investment function to back out—and thus control for—productivity. Building on OP, LP suggest the use of intermediate input demand in place of investment demand as a proxy for unobserved productivity. See Eberhardt and Helmers (2010) for an in-depth discussion of these 'structural' estimators. I use the opreg Stata routine developed by Yasar, Raciborski and Poi (2008) and I use the *levpet* Stata routine developed by Petrin, Poi and Levinsohn (2004), respectively. H_{jst} is treated as a freely variable input. I do not observe investment, and hence for Column 2 of Table 2 I derived a proxy variable in t as the difference between the reported book value of capital at time $t+1$ and its value in t. The way I constructed the proxy variable somehow exacerbates the measurement error problems typically associated with the proxy variable approach. In addition, augmenting my specification with this proxy variable reduces my sample size substantially, as (a) many firm-year observations are lost when I take the difference in reported book values and (b) the OP approach requires positive values for the proxy variable, eliminating additional firm-year observations. (The estimation routine will truncate firms' non-positive proxy variable observations because the monotonicity condition necessary to invert the investment function, and hence back out productivity, does not hold for these observations.)

A.VI Non-Tradable Goods

In Subsection 5.2 I used a dummy taking value one if the industry produces goods that are not widely traded outside the LLM. Industries for which the dummy takes value one are those classified as SMSA industries by Weiss (1974): Bottled and Canned Soft Drinks and Carbonated, Mineral, and Plain Waters; Fluid Milk; Bread and Other Bakery Products, Except Cookies and Crackers; Manufactured Ice; Primary Forest Products; Newspapers; Commercial Printing (except Lithographic); Commercial Printing (Lithographic); Engraving and Plate Printing; Typesetting; Photo-Engraving; Electrotyping and Stereotyping; Ready-Mix Concrete.

A.VII Back-of-the-envelope calculations details

Table 1 in GHM reports statistics for the sample of plants whose opening is considered in their study. These plants are quite large: they are more than twice the size of the average incumbent plant and account for roughly nine percent of the average county's total output one year prior to their opening. The mean output (five years after their assigned opening date) is 452,801, 000 of year-2006 dollars, or 395,476,000 of 2000 euros. Standard deviation is 901,690, 000 of year-2006 dollars. As explained in the notes of Table 1 in GHM, these statistics are for a subset of the 47 plant openings studied by the authors. In particular, a few very large outlier plants were dropped so that the mean would be more representative of the entire distribution (those dropped had output greater than half of their county's previous output and sometimes much more).

In order to establish the increase in the number of HWFs that a Veneto locality must experience to observe a change in local output comparable to the output increase caused by the opening of one large plant in GHM, I need to obtain the value of output for a typical HWF. Instead of dropping very large outlier plans as in GHM, I take the median of the distribution. The median value of output for HWFs in my sample is 7110 thousand of year-2000 euros. Therefore a Veneto locality must experience an increase of 395,476,000/7,110,000=56 HWFs. This is the change in my Back-of-the-envelope calculations.

A.VIII Additional References

Dal Bo E., F. Finan and A. Rossi, 2013 "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service", Forthcoming, Quarterly Journal of Economics.

Manning, A. and B. Petrongolo, 2013 "How local are labor markets? Evidence from a spatial job search model", Mimeo, LSE, http://personal.lse.ac.uk/petrongo/.

Ouazad A., 2007 "Program for the Estimation of Two-Way Fixed Effects", Mimeo, LSE, http://personal.lse.ac.uk/ouazad/.

Weiss, L., 1972. "The Geographic Size of Markets in Manufacturing." Rev. Econ. and Statis. 54 (August): 245–57.

A.IX Additional Figure and Tables

Figure A.1: *Mean wages of job changers within the same quantile of the AKM firm effect - all transition, all years (1992-2001)*

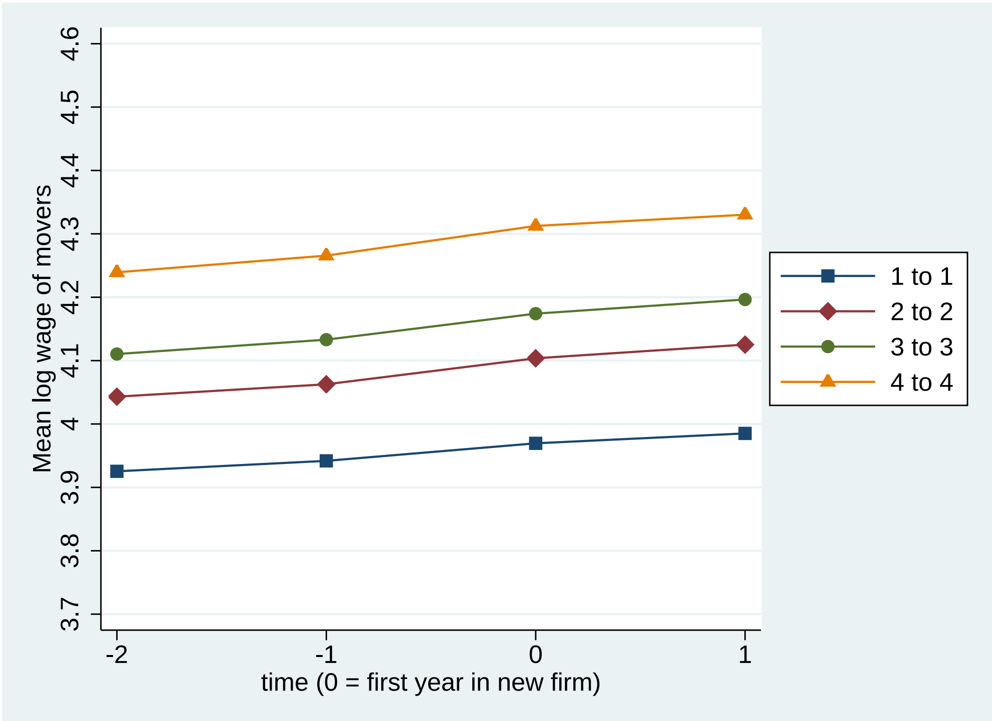


Figure A.2: *Mean wages of job changers from the 1st and the 4th quantile of the AKM firm effect - all transition, all years (1992-2001)*

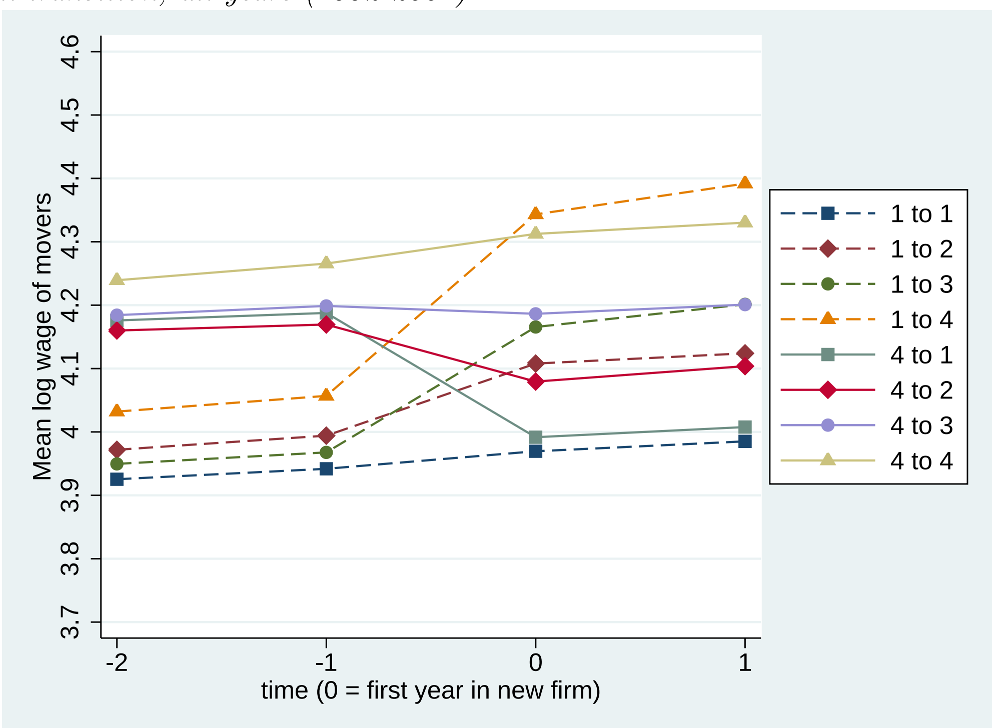
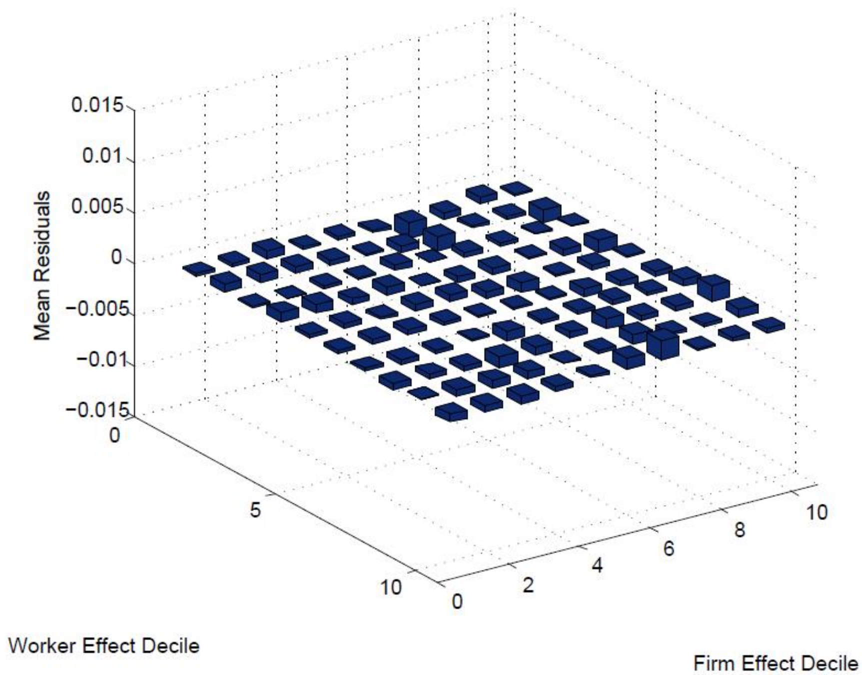
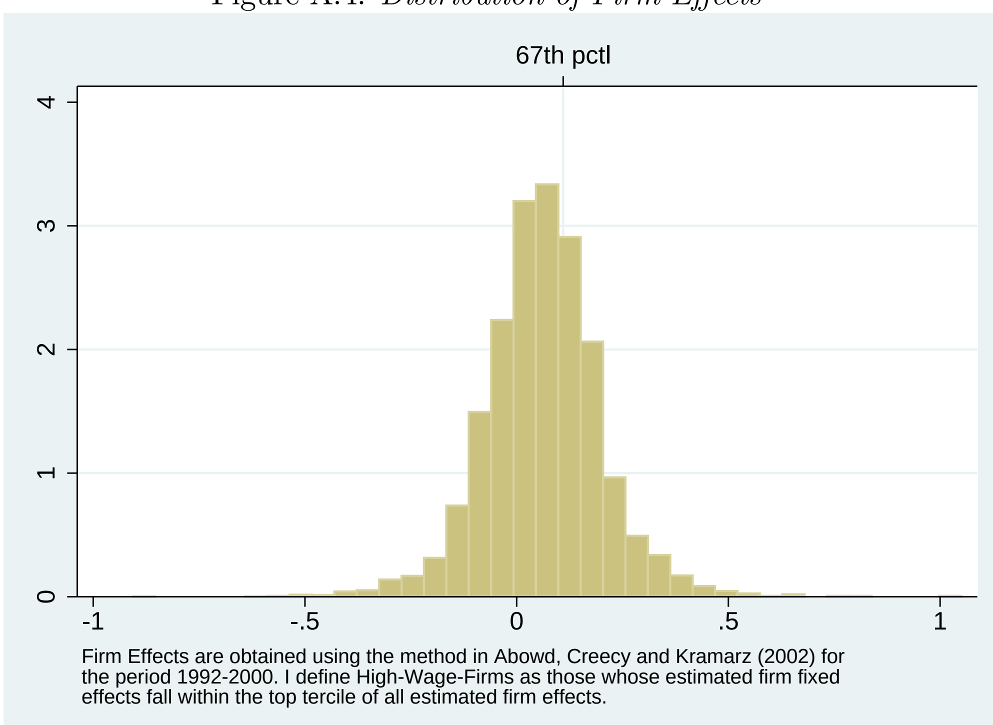


Figure A.3: *Mean Residuals by Person/Firm Effect Deciles*



Note: The figure shows mean residuals from the AKM regression by cells defined by decile of the estimated worker effect x decile of the estimated firm effect.

Figure A.4: *Distribution of Firm Effects*



{

Figure A.5: *Distribution of HWFs across Local Labor Markets (LLMs)*

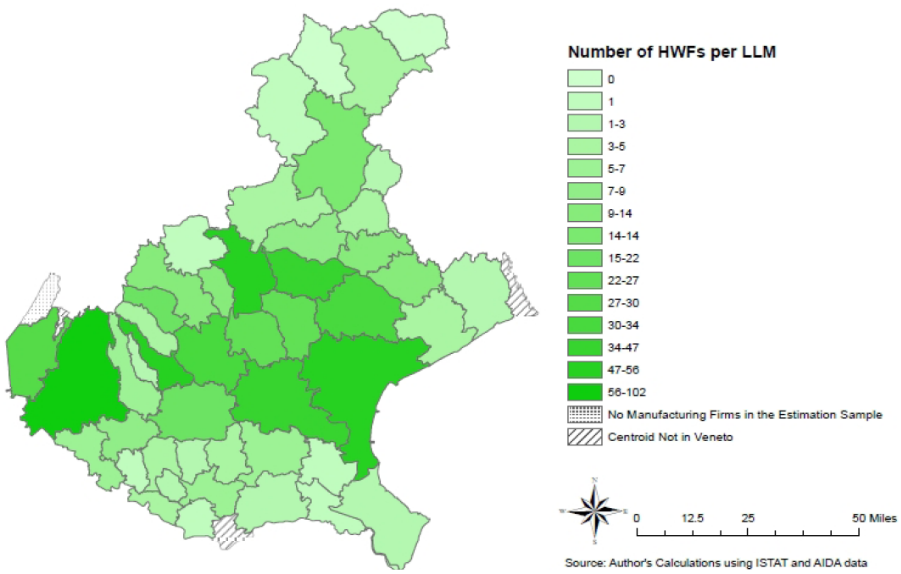


Table A.1: HWFs, Descriptive Statistics

Variable	Mean	(Std. Dev.)	Min.	Max.	N
Output	17170.817	(42613.033)	4.963	1354629.375	8996
Capital	4234.685	(12166.078)	0	356629.063	8984
Materials	9226.968	(24772.297)	0	739299.375	8996
Value added	4241.419	(12669.38)	-5835.655	475541.125	8984
Tangigle capital	3722.029	(10333.812)	0	195677.891	8984
Intangible capital	512.656	(3860.283)	0	182082.422	8984
firm age (years)	16.992	(11.068)	0	93	11007
employees from AIDA	80.288	(359.158)	1	20948	10988
employees from VWH	67.86	(145.494)	11	4896	13643
blue collars	38.939	(89.438)	0	3915	13643
white collars	14.865	(46.086)	0	1534	13643
managers	1.482	(8.427)	0	408	13643
female employees	14.277	(54.421)	0	2692	13643
employees age$<$ 30	17.73	(41.395)	0	1616	13643
employees age$>$ 45	13.201	(31.794)	0	795	13643
Downsize	0.178	(0.383)	0	1	12485
Larger Downsize	0.164	(0.37)	0	1	12485

Sample includes 1887 Individual Firms in the period 1995-2001. Output, Capital, Materials, Value Added are in thousands of 2000 euros. Employees from AIDA refers to the values found in the AIDA balance sheet data. Employees from VWH refers to the values obtained from head count in the Veneto Worker History data from Social Security. The variable Downsize takes value 1 if the drop in the labor force is larger than 1 percent, and the decrease in the labor force is greater than or equal to three individuals. The variable Larger Downsize takes value 1 if the drop in the labor force is larger than 5 percent, and the decrease in the labor force is greater than or equal to three individuals.

Table A.2: Characteristics of HWFs Workforce, 1992-2001

	(1) share white coll.	(2) share manager	(3) share female	(4) share age<30	(5) share age>45
HWF	0.022	0.004	-0.034	-0.003	-0.003
	(0.004)	(0.001)	(0.005)	(0.005)	(0.004)
Observations	42845	42845	42845	42845	42845
Adj. R-squared	0.214	0.112	0.556	0.156	0.126

All OLS regressions include year and 4-digit industry dummies. Standard errors (in parentheses) clustered by firm. The dummy HWF takes value 1 if the firm is classified as high-wage after estimating the AKM model on the period 1992-2000.

Table A.3: Job changes by direction of mobility

	All	Within LLM	Across LLM	Within ind.	Across ind.
HWF to HWF	12,461	8,112	3,999	6,980	5,481
HWF to non-HWF	7,732	4,097	3,398	2,688	5,044
non-HWF to HWF	12,831	7,065	5,501	5,087	7,744
non-HWF to non-HWF	28,709	18,175	10,011	12,395	16,314
Total	61,733	37,449	22,909	27,150	34,583

Table A.4: Share of workers in non-HWFs with HWF experience

Year	Experience from HWFs	Total workers with HWF experience	Total workers
1995	0.54	781	143,214
1996	0.70	1,058	150,421
1997	0.66	1,017	152,634
1998	0.73	1,124	153,395
1999	0.72	1,110	153,740
2000	0.81	1,251	154,456
2001	1.02	1,550	151,351

Table A.5: Share of non-HWFs employing workers with HWF experience

Year	% share of non-HWFs with H>0
1995	17.8
1996	21.4
1997	18.6
1998	19.6
1999	22.3
2000	23.4
2001	29.0

Table A.6: Movers from HWFs to non-HWFS, by occupation

Occupation	Number of workers
Apprentice	64
Blue collar	6,587
White collar	2,335
Manager	388
Total	9,376

Table A.7: Characteristics of knowledgeable vs. non-knowledgeable workers in non-HWFs.

	Knowledgeable workers			Non-knowledgeable workers			T-test of diff. of means	
Variable	Mean	(Std. Dev.)	N	Mean	(Std. Dev.)	N	Diff.	(Std. Err.)
age	34.227	(8.17)	678	35.659	(9.961)	142,536	-1.432***	(0.383)
female	0.215	(0.411)	678	0.326	(0.469)	142,536	-0.110***	(0.018)
blue collar	0.671	(0.47)	678	0.727	(0.446)	142,494	-0.055***	(0.017)
white collar	0.301	(0.459)	678	0.243	(0.429)	142,494	0.058***	(0.016)
manager	0.022	(0.147)	678	0.013	(0.113)	142,494	0.009**	(0.004)

Workers observed in 1995. Tables for all other years are very similar.

Table A.8: non-HWFs, Main Estimation Sample

Variable	**Mean**	**(Std. Dev.)**	**Min.**	**Max.**	**N**
Output	8536.626	(9745.812)	1101.159	94712.109	17158
Capital	1885.771	(2482.545)	58.229	21254.24	17158
materials	4328.5	(5795.519)	85.084	52073.469	17158
value added	2163.3	(2418.32)	-4082.134	36787	17158
Tangible Capital	1737.3	(2336.372)	2.833	20677.465	17158
Intangible Capital	148.471	(437.372)	0	12792.205	17158
firm age (years)	17.843	(10.864)	0	117	17158
employees from AIDA	49.717	(49.68)	2	450	17158
employees from VWH	50.835	(47.996)	11	482	17158
blue collars	31.079	(31.419)	0	348	17158
white collars	10.076	(13.123)	0	253	17158
managers	0.682	(1.881)	0	54	17158
female employees	13.572	(19.61)	0	309	17158
employees age$<$ 30	14.59	(14.48)	0	201	17158
employees age$>$ 45	9.348	(13.285)	0	199	17158
H workers	0.305	(0.727)	0	16	17158
H from same Ind	0.093	(0.384)	0	13	17158
H from diff Ind	0.212	(0.586)	0	16	17158
Exposure	0.05	(0.168)	0	5.602	17158

Table A.9: Local Variables: Summary Statistics, 1995-2001

Variable	Mean	(Std. Dev.)	Min.	Max.	N
lag (downsiz. HWFs)	1.026	(1.553)	0	11	17158
lag (downsiz. HWFs, larger drop)	0.937	(1.433)	0	10	17158
Local HWFs in same Ind	2.078	(6.236)	0	54	17158
Local non-HWFs in same Ind	6.854	(10.382)	1	54	17158
Lag 5 (Local HWFs in same Ind)	2.547	(4.92)	0	39	17158
Lag 5 (Local non-HWFs in same Ind)	9.133	(11.724)	1	50	17158

Table A.10: Knowledgeable Workers and Productivity in non-HWFs: additional specifications addressing endogenity concerns

	(1) Experience HWFs/LWFs	(2) Within	(3) Workforce Characteristics	(4) Mat-Cap-Lab Interactions t,t+1
log(capital)	0.097	0.065	0.091	
	(0.007)	(0.005)	(0.005)	
log(materials)	0.585	0.596	0.573	
	(0.011)	(0.013)	(0.007)	
log(employees)	0.224	0.060	0.229	
	(0.009)	(0.004)	(0.006)	
H workers	0.022	0.010	0.029	0.012
	(0.004)	(0.002)	(0.003)	(0.003)
Recent LWF exp	0.003			
	(0.002)			
$\beta_{\widetilde{H}}^{HWF} = \beta_{\widetilde{N}}^{LWF}$,pv	0.000			
Observations	8791	17158	17158	13540
Adj. R-squared	0.938	0.986	0.933	0.961

Dependent variable: Log(Output). Standard errors (in parentheses) clustered by firm. Regressions include industry-year interaction dummies and LLM-year interaction dummies. The variable 'H workers' is the number of knowledgeable workers currently observed at non-HWFs. Column 1 is estimated on the sample of Medium-Wage-Firms (MWFs) and includes workers with recent experience at HWF and Low-Wage-Firms (LWFs). '$\beta_{\widetilde{H}}^{HWF} = \beta_{\widetilde{N}}^{LWF}$,pv' is the p-value of the equality of coefficients of the variable 'Recent HWF exp' and the variable 'Recent LWF exp'. Column 2 reports within estimates. Column 3 adds the shares of managers, white collars, blue collars, females, and differently aged workers. Column 4 includes polynomial functions of capital, materials and number of employees in both t and t+1. This specification is in the spirit of the Ackerberg, Caves, and Frazer (2008) approach.

Table A.11: Knowledgeable Workers and Productivity in non-HWFs, 1995-2001, alternative grouping of firms based on TFP

	(1) OLS	(2) OP	(3) LP	(4) Inv-Cap Interactions t,t+1	(5) Mat-Cap Interactions t,t+1
log(capital)	0.071	0.051	0.139		
	(0.003)	(0.014)	(0.010)		
log(materials)	0.638	0.630		0.644	
	(0.003)	(0.004)		(0.007)	
log(employees)	0.142	0.125	0.136	0.112	0.128
	(0.004)	(0.006)	(0.003)	(0.008)	(0.004)
H workers	0.019	0.022	0.023	0.021	0.017
	(0.002)	(0.002)	(0.002)	(0.004)	(0.002)
Observations	16377	6225	16377	2712	12760
Adj. R-squared	0.940			0.942	0.945

I identify potential good firms as high-TFP firms. Specifically, I estimate firm effects from a specification in which the dependent variable is output, and I control for inputs. I identify good firms as those whose estimated firm fixed effects fall within the top third of all estimated firm effects. In this Table, H workers is the number of workers with experience at good (i.e. high-TFP) firms currently observed at non-good (non-high-TFP) firms.
Dependent variable: Log(Output). Standard errors (in parentheses) clustered by firm. Column 1 reports estimates using OLS. Column 2 implements the procedure in Olley and Pakes (1996). Column 3 implements the procedure in Levinsohn and Petrin (2003). Column 4 adds a third-degree polynomial function of log capital and log investment and the interaction of both functions in t and t+1. Column 5 includes the same controls as col. 5 but replaces log investment with log materials.

Table A.12: Knowledgeable Workers and Productivity in non-HWFs, Robustness to Different Specifications

	(1) Value Added	(2) Share	(3) Log	(4) Input-Industry Interactions	(5) Translog Functional Form
log(capital)	0.256	0.094	0.092	0.016	0.225
	(0.007)	(0.005)	(0.005)	(0.028)	(0.028)
log(employees)	0.551	0.228	0.223	-0.064	0.378
	(0.010)	(0.006)	(0.006)	(0.020)	(0.033)
H workers	0.079			0.026	0.009
	(0.005)			(0.003)	(0.003)
log(materials)		0.585	0.583	0.255	-0.306
		(0.007)	(0.007)	(0.055)	(0.035)
share of H workers		0.422			
		(0.131)			
log(H workers)			0.057		
			(0.010)		
No H workers			-0.034		
			(0.006)		
Observations	17116	17158	17158	17158	17158
Adj. R-squared	0.775	0.931	0.931	0.936	0.956

Log(Output) is the dependent variable in all columns excepts Column 1. In Column 1, Log(Value Added) is the dependent variable. Standard errors (in parentheses) clustered by firm. Regressions include industry-year interaction dummies and LLM-year interaction dummies. The variable 'H workers' is the number of knowledgeable workers currently observed at non-HWFs. The variable 'log(H workers)' is the logarithm of number of knowledgeable workers. The dummy 'No H workers' takes value 1 if the number of knowledgeable workers is equal to 0. Column 1 reports estimates with Log(Value Added) as dependent variable. Column 2 replaces the number of H workers with the share of H workers. Column 3 replaces the number of H workers with the log of H workers plus the dummy 'No H workers'. Column 4 allows the effect of each input to differ by two-digit industry level. Column 5 uses a translog functional form for inputs.

Table A.13: Knowledgeable Workers and Productivity in non-HWFs, Further Extensions, 1995-2001

	(1) Same/Diff Industry	(2) Previous Occupation	(3) Current Occupation	(4) Continuous Measure	(5) Lag
H from same Ind	0.035				
	(0.006)				
H from diff Ind	0.028				
	(0.004)				
H current higher-skilled occ.		0.042			
		(0.006)			
H current lower-skilled occ.		0.025			
		(0.004)			
H previous higher-skilled occ.			0.044		
			(0.006)		
H previous lower-skilled occ.			0.025		
			(0.004)		
Exposure				0.026	
				(0.013)	
Lag (H Workers)					0.024
					(0.005)
$\beta_H^{same} = \beta_H^{diff}$,pv	0.232				
$\beta_H^{high} = \beta_H^{low}$,pv		0.013	0.006		
Observations	17158	17158	17158	17158	16265
Adj. R-squared	0.931	0.931	0.931	0.931	0.932

Dependent variable: Log(Output). All columns include log(capital), log(labor) and log(employees). Standard errors (in parentheses) clustered by firm. Column 1 differentiates between workers moving within the same industry and between industries. $\beta_H^{same} = \beta_H^{diff}$,pv is the p-value of the equality of coefficients of the variable 'H from same Ind' and the variable 'H from diff Ind'. In Column 2 H is disaggregated into two groups based on the occupation at the previous employer (HWF). In Column 3 it is disaggregated based on the occupation at the current employer (non-HWF). $\beta_H^{high} = \beta_H^{low}$,pv is the p-value of the equality of coefficients of the H workers in higher-skilled occupations and the H workers in lower-skilled occupations. In Column 4 I employ an alternative, continuous, measure of the receiving firm's exposure to knowledge, which exploits the differences between sending and receiving firms, thus extending the analysis above which has so far worked with a dummy indicating experience at a HWF. (see text for details on the definition of this variable) In Column 5 the variable of interest is lagged by one year.

Table A.14: Knowledgeable Workers with experience in the same industry and Productivity in non-HWFs, 1995-2001

	(1) OLS	(2) OP	(3) LP	(4) Inv-Cap Interactions	(5) Inv-Mat Interactions	(6) IV
H from same Ind	0.036	0.037	0.031	0.022	0.026	0.121
	(0.006)	(0.008)	(0.005)	(0.012)	(0.006)	(0.062)
Observations	17158	6635	17158	2963	13540	17158
Adj. R-squared	0.931			0.940	0.952	0.921
Fstat, instrum., 1st stage						27.22

Dependent variable: Log(Output). All columns include log(capital), log(labor) and log(employees). Standard errors (in parentheses) clustered by firm. The variable 'H workers same Ind' is the number of workers with HWF experience in the same industry currently observed at non-HWFs. Column 1 reports estimates from the baseline specification. Column 2 implements the procedure in Olley and Pakes (1996). Column 3 implements the procedure in Levinsohn and Petrin (2003). Column 4 adds a third-degree polynomial function of log capital and log investment and the interaction of both functions in t and t+1. Column 5 includes the same controls as col. 5 but replaces log investment with log materials. Column 6 reports IV estimates using the lagged number of downsizing local good firms in the same 4-digit industry. It includes an indicator of the importance of local demand, namely a dummy taking value 1 if the 4-digit industry produces goods that are not widely traded outside the LLM. It also controls for an index of industry localization, namely the ratio between the number of firms in Veneto and total Italian firms in a given 4-digit industry.

Table A.15: Number of local HWFs and Productivity, Estimates, 1995-2001

	(1) OLS	(2) System GMM	(3) Two-step System GMM
log(capital)	0.0940	0.0176	0.0250
	(0.0050)	(0.0586)	(0.0695)
log(materials)	0.5849	0.6275	0.6215
	(0.0109)	(0.0463)	(0.0555)
log(employees)	0.2295	0.0034	0.0044
	(0.0104)	(0.0109)	(0.0117)
Lag 5 (Local HWFs in same Ind)	0.0028		
	(0.0007)		
Lag 5 (Local non-HWFs in same Ind)	-0.0008		
	(0.0003)		
l.log(output)		0.9839	0.9766
		(0.0545)	(0.0559)
l.log(capital)		-0.0274	-0.0320
		(0.0527)	(0.0650)
l.log(materials)		-0.6140	-0.6077
		(0.0502)	(0.0573)
l.log(employees)		0.0095	0.0118
		(0.0154)	(0.0184)
Local HWFs in same Ind		0.0009	0.0010
		(0.0006)	(0.0006)
Local non-HWFs in same Ind		-0.0010	-0.0010
		(0.0006)	(0.0006)
$\beta_{HWFs} = \beta_{non-HWFs}$,pv	0.000	0.076	0.072
Observations	17158	13501	13501
AR(1)z		-11.88	-10.60
AR(2)z		0.782	0.827
AR(3)z		1.954	1.969
AR(4)z		-1.147	-1.169
HansPv		0.874	0.874
Adj. R-squared	0.931		

Dependent variable: Log(Output). Standard errors (in parentheses) clustered by LLM. Regressions include year dummies. Column 1 reports OLS estimates. Column 2 reports System GMM estimates. Column 3 reports twp-step System GMM estimates, using Windmeijer-corrected standard errors. In Column 2 and 3 the variables 'Local HWFs in same industry' and 'Local non-HWFs in same industry' are treated as endogenous. AR(1)z, AR(2)z, AR(3)z, AR(4)z: Arelanno and Bond (1999) test of first, second, third and fourth order serial correlation, distributed as N(0,1). HansPv: p-value of Hansen test of overidentifying restrictions. Only the shortest allowable lagged is used as instrument. $\beta_{HWFs} = \beta_{non-HWFs}$,pv is the p-value of the equality of coefficients of the variable 'Local HWFs in same industry' and the variable 'Local non-HWFs in same industry'.