"Good" Firms, Worker Flows, and Local Productivity

Michel Serafinelli, *University of Essex and Institute for the Study of Labor (IZA)*

This paper is the first to present direct evidence showing how localized knowledge spillovers arise from workers changing jobs within the same local labor market. Using a unique data set combining Social Security earnings records and balance sheet information for the Veneto region of Italy, I first identify a set of highly productive firms, then show that hiring workers with experience at these firms significantly increases the productivity of other firms. My findings imply that worker flows explain around 10% of the productivity gains experienced by incumbent firms when new highly productive firms are added to a local labor market.

I. Introduction

A prominent feature of the economic landscape in many developed countries is the tendency for firms to locate near other firms producing similar

I am very grateful to David Card, Enrico Moretti, Patrick Kline, and Bronwyn Hall for invaluable guidance. I also thank Miguel Almunia, Vladimir Asriyan, Pamela Campa, Federico Callegari, Jeff Chan, Luca Citino, Kunal Dasgupta, Francesco Devicienti, Sabrina Di Addario, Sebastian Findeisen, Luisa Gagliardi, Carlo Gianelle, Albrecht Glitz, Jonas Hjort, Mitchell Hoffman, Shimeng Liu, Agata Maida, Robert McMillan, Peter Morrow, Marco Percoco, Steven Raphael, Ana Rocca, Fabio Schiantarelli, Fabiano Schivardi, Aloysius Siow, Grigorios Spanos, Bastian Stockinger, Matthew Turner, Marco Valentini, Victoria Vanasco, Alex Whalley, and Yves Zenou as well as seminar participants at the National Bureau of Economic Research Summer Institute, the American Economic Association, Wharton, Universitat Pompeu Fabra, Warwick, the Paris School of Economics, the Society of Labor Economists, Berkeley, the University of Calgary, the University of Cambridge, Stockholm University,

[*Journal of Labor Economics*, 2019, vol. 37, no. 3]
 0734-306X/2019/3703-0004$10.00
Submitted March 24, 2016; Accepted April 4, 2018; Electronically published April 25, 2019

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. Furthermore, the growth and diffusion of multinational corporations has led to the recent appearance of important industrial clusters in several emerging economies (Alfaro and Chen 2014).

Researchers have long speculated that firms in such 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 that significant productivity advantages of agglomeration exist for many industries (Henderson 2003; Rosenthal and Strange 2003; Cingano and Schivardi 2004; Ellison, Glaeser, and Kerr 2010; Greenstone, Hornbeck, and Moretti 2010; Combes et al. 2012; Gathmann, Helm, and Schönberg 2016). Disagreement remains, however, over the nature of the microeconomic mechanisms that can account for these advantages (Glaeser and Gottlieb 2009; Moretti 2011). This serves as a barrier to understanding differences in productivity across industry clusters and localities and hinders the design of location-based policies (Glaeser and Gottlieb 2008; Kline 2010).

Localized knowledge spillovers are one of the most commonly hypothesized sources of the productivity advantages of agglomeration, alongside the availability of specialized intermediate inputs, the sharing of a common labor pool, and better matching. Nevertheless, if information can easily flow out of firms, the question why the effects of spillovers are localized must be clarified—a point well made by Combes and Duranton (2006).

This is the first paper to present direct evidence showing how firm-to-firm labor mobility enhances the productivity of firms located near highly productive firms. In doing so, it lends support to the idea that the strong localized aspect of knowledge spillovers discussed in the agglomeration literature arises—at least in part—from the propensity of workers to change

Universitat Autònoma de Barcelona, the University of Alicante, Collegio Carlo Alberto, the University of Mannheim, Fondazione Rodolfo Debenedetti, the Einaudi Institute for Economics and Finance, Universidad Torcuato Di Tella, Urban Economic Association meeting, and the University of Toronto for comments. I acknowledge financial support from the Berkeley Center for Equitable Growth. Thanks are due also to Veneto firms' employees who have generously given their time in interviews and to Ishita Arora, Jessica Burley, Olivia Casey, Benjamin Couillard, Helen Kissel, Nitin Kohli, and Palermo Penano for excellent research assistance. I am indebted to Giuseppe Tattara for making available the Veneto Workers History data and to the University of Toronto Connaught Fund for financial support. Contact the author at michel.serafinelli@essex.ac.uk. Information concerning access to the data used in this paper is available as supplemental material online

jobs within the same local labor market (LLM): knowledge is partly embedded in workers and diffuses when workers move between firms.[1]

To fix ideas, I begin by presenting a simple conceptual framework in which some firms are more productive because they possess superior knowledge. The superior knowledge could include information about export markets, physical capital, new organizational forms, and intermediate inputs. Employees at these firms acquire relevant firm's internal knowledge—for simplicity, I refer to these employees as "knowledgeable" workers. Other firms can then gain access to this superior knowledge by hiring such workers.

The central empirical goal of the paper is to measure the importance of labor market–based knowledge spillovers. In confronting the nontrivial measurement challenges involved, I take advantage of a unique data set that combines Social Security earnings records and detailed financial information for firms in Veneto, a region of Italy with many successful industrial clusters. In the empirical analysis, 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. For convenience, I refer to these HWFs as "good" firms.

Next, I evaluate the extent to which other firms benefit from hiring knowledgeable workers by studying the effect on productivity associated with hiring workers with recent experience at these good firms. Given that workers do not move from firm to firm on a random basis, my analysis addresses important identification threats. In particular, positive productivity shocks that are correlated with the propensity to hire knowledgeable workers may give rise to an upward bias in the estimated impact of such knowledgeable workers.[3] To address this potential endogeneity issue, I use well-established control function methods drawn from the productivity literature (Olley and Pakes 1996; Levinsohn and Petrin 2003).[4] In doing so, I find that, on average, recruiting a knowledgeable worker increases the productivity of

[1] Other possibilities include various forms of communication externalities: face-to-face meetings, word-of-mouth communication, and direct interactions between skilled workers from different firms (Charlot and Duranton 2004).

[2] This 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. Results are similar when using alternative groupings of firms based on output (controlling for inputs) and value added.

[3] Examples of such shocks are process innovations and new managerial techniques.

[4] Olley and Pakes (1996) 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 Olley and Pakes

a non-high-wage firm (non-HWF) by between 1.8% and 3%.[5] Additional evidence supports the main finding that the recruitment of workers with experience at good firms significantly increases the productivity of the non-HWFs hiring them. I observe greater productivity gains in firms hiring workers in higher-skilled occupations. The productivity effect of knowledgeable workers does not appear to be driven by unobserved worker quality,[6] and it is not associated with recently hired workers in general; placebo regressions show that there is no similar productivity effect for recently hired workers with experience at firms that have lower productivity than the receiving firm. These results indicate that the estimated effect is not due to better worker-firm matching or to switchers being more productive than stayers in general (regardless of previous employment history). I also rule out the possibility that the results are driven by time-invariant unobservables, such as managerial talent.

It is also possible that knowledgeable workers are attracted to join firms that are "on the rise," (i.e., firms that offer better prospects than the good firms at which these workers are employed) rather than knowledgeable workers moving to firms and causing the increase in productivity. To address this concern, I adapt control function methods to proxy for future productivity shocks. In addition, I instrument for the number of knowledgeable workers employed by a non-HWF with the number of good firms locally in the same industry that downsized in the previous year. Following a downsizing event at a good firm, it is more likely that a knowledgeable worker applies for a job at local non-HWFs because she is unemployed and does not want to relocate far away. Put differently, in the scenario captured by the instrumental variable (IV) approach, the strategic mobility explanation is less likely to play a role.[7] Applying this approach, the IV esti-

(1996), Levinsohn and Petrin (2003) suggest the use of intermediate input demand in place of investment demand as a proxy for unobserved productivity.

[5] In interpreting my estimates, it is important to highlight that non-HWFs are quite small: their median number of employees is 33. Furthermore, as many as 78% of non-HWFs in a given year do not employ any knowledgeable workers. Hiring one knowledgeable worker therefore implies a significant change in terms of workforce for most firms in my data.

[6] To investigate this issue, I obtain a proxy for worker ability using estimates of worker fixed effects from wage equations where both firm and worker effects can be identified.

[7] While the timing of these moves is arguably exogenous, knowledgeable workers may still decide which new employer to join among the set of non-HWFs after being displaced by good firms. However, in small labor markets and specialized industries, workers are likely to have a limited set of alternatives. Note also that this is a new approach: the number of downsizing local good firms is used as an instrument for a firm input in a production function framework, and to date only past values of the regressors themselves or input prices have been used for instrumentation in the production function literature (see the survey by Eberhardt and Helmers 2010).

mates return an economically and statistically significant coefficient on the number of knowledgeable workers. The exclusion restriction would be violated and the coefficient on knowledgeable workers biased upward if there were localized unobservable industry-specific shocks that led good firms to downsize and positively affect measured productivity at non-HWFs. An example of such shocks is unobserved shifts in local demand from the products of HWFs to the products of non-HWFs. I obtain data on the level of tradability of goods and show reassuring evidence regarding the relevance of such shocks.

In the second 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. Here, I relate my findings on the effect of firm-to-firm labor mobility to the existing evidence on the productivity advantages of agglomeration, focusing on a study by Greenstone, Hornbeck, and Moretti (2010). These authors find that following the opening of a large manufacturing plant, the TFP of incumbent plants in the US counties that were able to attract these large plants increased 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 in Greenstone, Hornbeck, and Moretti (2010) is larger if incumbent plants are "economically" close to the large plant; measures of economic links include a dummy indicating belonging to the same industry and indicators of technological linkages and worker flows at the industry level. Furthermore, this TFP effect increases over time. These facts are consistent with the presence of intellectual externalities that are embodied in workers who move from firm to firm. However, data limitations prevented Greenstone, Hornbeck, and Moretti (2010) from drawing definitive conclusions regarding the underlying mechanism. Building on their analysis, I am able to evaluate the extent to which worker flows explain the productivity advantages of agglomeration by predicting the change in local productivity following an event analogous to that studied by Greenstone, Hornbeck, and Moretti (2010) within the worker mobility framework described above. I find that the predicted change in productivity is around 10% of the overall local productivity change observed after the event.[8] Finally, I show that the local productivity effect attributed to good firms does not appear to be associated with a general increase

[8] The remaining portion is likely to be explained by other types of knowledge spillovers not based on labor mobility (i.e., various forms of communication externalities) and the availability of specialized intermediate inputs. That labor market–based knowledge spillovers account for 10% of agglomeration advantages is consistent with the estimates in Ellison, Glaeser, and Kerr (2010), who find that knowledge spillovers do not account for a large fraction of agglomeration economics, and with the estimates in Kantor and Whalley (2014), who find smaller spillovers from universities than Greenstone, Hornbeck, and Moretti (2010) find from large manufacturing plants.

in the size of the labor market: large productivity gains linked to changes in the number of firms seem to be realized only when the new entrants are good firms. This evidence suggests that the estimated impact does not reflect better worker-firm matching arising from a thicker labor market.

The remainder of this paper is organized as follows. In Section II I relate my research to the existing literature. Section III presents a conceptual framework that guides the empirical exercise and then discusses my econometric strategy. In Section IV I describe the data and present descriptive results. The main regression results, in addition to various extensions and robustness checks, are presented in Sections V, VI, and VII. Section VIII concludes.

II. Relation to Previous Research

This paper adds to the growing literature examining productivity advantages through agglomeration, a literature reviewed in Duranton and Puga (2004), Rosenthal and Strange (2004), and Moretti (2011). The research relating most closely to this paper studies the micro foundations of agglomeration advantages based on knowledge spillovers. In the theoretical analysis by Combes and Duranton (2006), 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).[9] In a case study of the British Motor Valley, Henry and Pinch (2000, 198–99) conclude the following:

> 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.

I contribute to the literature on the micro foundations of agglomeration advantages by showing direct evidence of productivity gains through worker flows.[10]

[9] The study of research-and-development 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.

[10] Given the Veneto industry mix, discussed in Sec. IV, our findings are more relevant for traditional manufacturing regions, such as Germany's Baden-Wuerttemberg and the British Motor Valley than for high-tech regions like Silicon Valley, analyzed by Saxenian (1994), or Cambridge, United Kingdom.

Some research outside the agglomeration literature has also emphasized 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 include the studies by Cooper (2001), Markusen (2001), Glass and Saggi (2002), and Fosfuri, Motta, and Rønde (2001). Several convincing firm-level empirical analyses have already been conducted: Poole (2013) finds a positive effect of the share of new workers previously employed by foreign-owned firms on wages paid in domestic firms in Brazil; Balsvik (2011) offers a detailed account of productivity gains linked to worker flows from foreign multinationals to domestic firms in Norway; Parrotta and Pozzoli (2012) provide evidence from Denmark regarding the positive impact of the recruitment of knowledge carriers (technicians and highly educated workers recruited from a donor firm) on a firm's value added; and Stoyanov and Zubanov (2012) show that Danish firms that hired workers from more productive firms become more productive.[11]

My findings are consistent with the recent empirical contributions of Poole (2013), Balsvik (2011), Parrotta and Pozzoli (2012), and Stoyanov and Zubanov (2012) that study worker flows using linked worker-firm data. In terms of the firm-level analysis, which is the focus of the first part of my paper, I build on their research and offer (1) a new definition of "knowledgeable" workers and (2) additional hypotheses testing that leaves less room for alternative interpretations of the main results. Unlike the above authors, who focus exclusively on the relationship between labor mobility and productivity at the firm level, I also seek to shed light on a broader question: To what extent can labor mobility explain the productivity advantages through agglomeration (at both firm and LLM levels)? While the issues analyzed in this paper are of general interest, the specific case of Veneto is also important. This region is part of a larger economic area in Italy where net-

[11] The related literature also includes Agrawal, Cockburn, and McHale (2006), who look at patent citations when workers move across firms; Irarrazabal, Moxnes, and Ulltveit-Moe (2013), who argue that worker heterogeneity accounts for much of the exporting premium of firms; and De la Roca and Puga (2017), who look at mobility of workers across cities. Another related body of work analyzes peer effects in the workplace induced by knowledge spillovers and finds mixed evidence. On one hand, Waldinger (2010) finds that faculty quality is a very important determinant of PhD student outcomes. On the other hand, Cornelissen, Dustmann, and Schönberg (2016) find only small peer effects in wages in high-skilled occupations, and Waldinger (2012) shows that even very high-quality scientists do not affect the productivity of their local peers. Guryan, Kroft, and Notowidigdo (2009) study teammates in golf and find no evidence of knowledge spillovers. Other papers within this body of work, focusing on social pressure, report productivity spillovers (Mas and Moretti 2009; Bandiera, Barankay, and Rasul 2010).

works of specialized firms, frequently organized in districts, have been effective in promoting and adapting to technological change during the last 3 decades. This so-called Third Italy region has received a good deal of attention by researchers, both in the United States and in Europe (Brusco 1983; Piore and Sabel 1984; Trigilia 1990; Piore 2009).

III. Framework and Empirical Strategy

A. Conceptual Framework

Consider a finite number of locations, each constituting a separate LLM. To fix ideas, assume that these labor markets are completely segmented, with workers being immobile among them. There exists a finite collection $\mathcal{J} = \{\mathcal{J}_0, \mathcal{J}_1\}$ of firms consisting of the set $\mathcal{J}_1$ of good firms, which are more productive because they have some superior knowledge, and the set $\mathcal{J}_0$ of other (nongood) firms, which have no access to 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 the relevant firm's internal knowledge.[12] Workers are either knowledgeable or unknowledgeable. All workers employed by good firms, because of their access to the firm's internal knowledge, are knowledgeable. Additionally, this knowledge can be transferred to a $j \in \mathcal{J}_0$ firm if the workers switch employers.

The production function of firm $j \in \mathcal{J}_0$ is

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

where $L_j = H_j + N_j$, that is, the sum of knowledgeable workers H_j (who moved at some point from a good firm to a nongood firm) and unknowledgeable workers N_j; K_j is total capital inputs; and M_j is material inputs. The managerial technology involves an element of diminishing returns to

[12] For simplicity, I assume that this acquisition of internal knowledge takes place immediately after the workers join the good firm. I also assume that this type of knowledge cannot all be patented and that exclusive labor contracts are not available. An interesting aspect, which is, however, not empirically investigated in this paper, concerns firm-level heterogeneity in the incentives to keep the knowledge secret and the strategy to prevent the transfer. In an email interview, an industry expert discussed this very issue: "Veneto firms for several reasons are often unwilling to patent, but they keep some trade secrets, at least as tacit knowledge. . . . Think of prosecco firms each producing a wine different from the other; this means that the added components, eg. sugar or yeast, and blends are part of the secret. . . . A key player in this regard is the oenologist who *de facto* establishes the 'wine formula.' . . . Trade secrets exist not only in large firms eg. Zonin or Ferrari, but also in cooperatives like Caviro, with the secret guarded by an aggregator who then governs the chain."

scale, or to "span of control" ($\delta < 1$).[13] I allow for knowledge transfer by letting efficiency depend on H_j:

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

where A_j denotes TFP and D_j denotes the component of TFP due to factors other than H. See appendix section A.I (the appendix is available online) for a formal discussion of the firm optimization problem.

B. Empirical Strategy

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

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

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. Section V.C reports estimation results for alternative specifications (in terms of functional forms of H and measures of productivity).[14] The variable of interest, H (the number of knowledgeable workers), is constructed from head counts in the matched employer-employee data.[15] I define a worker as being knowledgeable (having recent experience at a good firm) in year t if he or she is observed working in a good firm for one or more of the years $t-3$ to $t-1$.[16] In what follows, I use "knowledgeable workers" and "workers with recent experience at good firms" interchangeably.

The structure of regression equation (3) is in line with that in Greenstone, Hornbeck, and Moretti (2010), who also regress firm-level output on inputs

[13] This is in line with the large presence, which I document below, of small and medium-sized firms in the sample of nongood firms. This description of management follows Lucas (1978, 51) who assumes that "each agent is endowed with a managerial talent level x, drawn from a fixed distribution $\Gamma : R^+ \to [0, 1]$. If agent x manages resources n and k, his 'firm' produces $xg[f(n, k)]$ units of output, where $g : R^+ \to R^+$ is twice-differentiable, increasing, and strictly concave, satisfying $g(0) = 0$. . . . Whatever managers do, some do it better than others. Given this assumption, however, one is led immediately to the question: why does the best manager not run everything? Therefore, I assume concavity in the function g."

[14] Notice that $\beta_L = \delta\alpha$, $\beta_K = \delta\gamma$, and $\beta_M = \delta\lambda$.

[15] In Sec. V.D 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 that works with a dummy indicating experience at a good firm. Furthermore, I present estimates when I lag the number of workers with experience at good firms.

[16] It may be instructive to consider a practical example. Consider a worker who separates from a good firm in 1994 and joins nongood firm 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. Results are similar if I consider a 5-year window instead of a 3-year one.

(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 VI, to relate directly my findings regarding the effect of firm-to-firm labor mobility to the evidence in Greenstone, Hornbeck, and Moretti (2010).

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 other than H. The time-variant ζ_{jslt} represents deviations from this mean efficiency level and captures (i) unobserved factors affecting firm output, (ii) measurement error in inputs and output, and (iii) 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). 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 V.C 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.

1. *Unobserved Productivity Shocks*

Unobservable Contemporaneous Shocks

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

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

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 (ω_{jslt}), 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

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

One major difficulty in estimating β_H in equation (5) is that nongood firms may decide on their choice of H based on the realized firm-specific productivity shock ω_{jslt} unknown to the researcher (for a discussion using the first-order condition with respect to H obtained from the optimization problem, see app. sec. A.I). To assess the relevance of this issue in my setting, I present in Section V.A estimates using control function methods drawn from the productivity literature (Olley and Pakes 1996; Levinsohn and Petrin 2003). The Olley and Pakes (1996) approach addresses the issue of endogeneity of inputs by inverting the investment function to back out—and thus control for—productivity. Specifically, the approach uses information about observed investment to proxy for unobserved productivity and applies a control function estimator. Olley and Pakes (1996) assume that capital and the productivity shocks are firm-specific state variables in the firm's dynamic programming problem. Investment is chosen at time t and adds to the capital stock at time $t + 1$. The solution of the Bellman equation gives an investment policy function that depends on capital and productivity. Labor inputs are not included in the investment equation because they are assumed to be nondynamic inputs: they can be adjusted after realization of the productivity shock within the same period. A key assumption is that investment is strictly increasing in both capital stock and productivity. In addition, the productivity shock is assumed to be the only unobservable driving the investment choice. Finally, when deciding on investment in period $t + 1$, any realizations of the productivity shock prior to time t are not incorporated in the investment function because productivity evolves by assumption following an exogenous first-order Markov process: a firm builds expectations about its productivity at time $t + 1$ exclusively based on its productivity levels realized at time t. Therefore, one can assume most generally that productivity evolves according to $\omega_{jslt} = g(\omega_{jslt-1}) + \xi_{jt}$, where ξ_{jt} is the random productivity shock. Provided the investment function is continuous in capital and ω_{jslt} and provided investment is positive, the investment equation can be inverted to yield $\omega_{jslt} = f_t[i_{jslt}, \ln(K_{jslt})]$, where i_{jslt} denotes firm j's investment in physical capital at time t. Olley and Pakes (1996) propose estimation based on a third-order polynomial series expansion to approximate $f_t[i_{jslt}, \ln(K_{jslt})]$. Building on Olley and Pakes (1996), Levinsohn and Petrin (2003) suggest the use of intermediate input demand instead of investment demand as a proxy for productivity. This means that the decision on intermediate input is made at time t once the productivity shock is observed by the firm. The same applies to labor input choices, which in turn means that labor and intermediate inputs are chosen at the same time and labor inputs preserve their assumed nondynamic/flexible nature. In the Levinsohn and Petrin (2003) approach, intermediate inputs (electricity, material inputs) are modeled as a function of the productivity shocks and capital similar to the use of investment in the Olley and Pakes (1996) procedure. See Eberhardt and Helmers (2010) for a more in-depth discussion of these structural estimators, and see Section V.A for details on the estimation using my data.

Unobservable Future Shocks: Using the Number of Downsizing Firms as the IV

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. Specifically, 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 address this concern in Section V.B, I present estimates instrumenting for the number of knowledgeable workers in a nongood firm with the number of local good firms in the same two-digit industry that downsized in the previous period. Following a downsizing event at a good firm, it is more likely that a knowledgeable worker applies for a job at local nongood firms because she is unemployed and does not want to relocate far away. Put differently, in the scenario captured by the IV approach, the strategic mobility explanation is less likely to play a 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. 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 lay off workers. Second, 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 nongood firms. However, in small labor markets and specialized industries, workers may have a limited set of alternatives.[17]

The choice of the instrument is based on the notion that geographic proximity plays an important role in determining worker mobility. The evidence from targeted interviews supports the idea that distance acts as a barrier for job mobility. Moreover, in Section IV.B I show descriptive evidence regarding the propensity of workers to change jobs within the same LLM. Notice that the number of downsizing local good firms is an external instrument used in a production function framework. As pointed out in the survey by 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.

[17] The structure of the social network when the layoff occurs might predict which firm a worker joins, but this research question is beyond the scope of the current paper.

Alternative Approach: Proxy for Future Shocks

To explore the possibility of future productivity shocks further, I adapt the Olley and Pakes (1996) and Levinsohn and Petrin (2003) approaches and control for (i) polynomial functions of capital and investment in both t and $t + 1$ and (ii) polynomial functions of capital and materials, again in both t and $t + 1$, 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 V.A I provide the estimates and a longer discussion of such an approach. In Section V.C I also provide estimates when including polynomial functions of capital, materials, and the number of employees in both t and $t + 1$. This specification is in the spirit of the approach of Ackerberg, Caves, and Frazer (2015).

2. *Estimation of the Wage Model and Identification of Good Firms*

Empirically, I identify potentially high-productivity firms as HWFs: firms that pay a relatively high firm-specific wage premium. The use of alternative groupings of firms based on output controlling for inputs and value added yields very similar results (see Sec. V.C). The definition of high-productivity firms as 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 TFP and value added than other firms in my sample. There are three reasons why, for the baseline results, I use Social Security data to define good firms as HWFs (rather than define the good firms directly as the high-TFP or high-value-added ones and detect them using balance sheet data). First, Social Security data are available for a longer period of time than the balance sheets, and therefore their use allows a more accurate categorization of firms. Second, since Social Security records are administrative data, measurement error is lower than in balance sheets. Third, Social Security data allow estimation of a worker-level wage equation, controlling for measured individual characteristics and worker effects. 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 (Sec. V.C).

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

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

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.[18] The presence of labor mobility in matched worker-firm data sets (like the one I use) 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. See appendix section A.III for a detailed discussion and the results of the analyses I perform to test for such sorting.

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. Section IV.A reports descriptive results as well as more details on the estimation procedure. In what follows, I use "HWFs" and "good firms" interchangeably. Results 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 V.D I remove the top-third threshold and employ a continuous measure of the receiving firm's exposure to knowledge. This alternative measure is the difference in quality between the sending and the 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. This procedure extends the baseline analysis, which works with a dummy indicating experience at a good firm.

IV. 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% 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.[19] Manufacturing firms in Veneto specialize in metal engineering, goldsmithing, plastics, furniture, garments, textiles, leather, and shoes.[20] The manufacture of food and beverage—wine and baked goods, in particular—is also a prominent subsector.

[18] 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.

[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.

My data set pools three sources of information: individual earnings records, firm balance sheets, and information on LLMs.[21] The earnings records come from the Veneto Workers History (VWH) data set. The VWH has data on all private sector personnel in the Veneto region. 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 Analisi Informatizzata delle Aziende (AIDA), a database circulated by Bureau Van Dijk containing official records of all incorporated nonfinancial Italian firms with annual revenues of at least €500,000. AIDA's balance sheets include each firm's 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 appendix section A.II.

Information on LLMs is obtained from the National Institute of Statistics. 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.

To bolster the analysis, in January 2012 I visited several Veneto firms and interviewed employees about the history of their enterprises and their current operations. In the following years I have also conducted targeted interviews (via phone and email) with academic and industry experts as well as officials of chambers of commerce.

A. Abowd-Kramarz-Margolis (AKM) Estimation and Descriptive Statistics

This section reports descriptive results for the firms in my sample as well as more details on the AKM estimation procedure. 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

[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).

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

Table 1
Characteristics of High-Wage Firms (HWFs)

	Individual Wage (1)	TFP (2)	Value Added (3)	K (4)	Intangible K (5)
HWF	.130	.080	.105	.101	.274
	(.003)	(.008)	(.010)	(.025)	(.043)
Observations	1,837,597	26,657	26,587	26,674	24,450
Adjusted R^2	.912	.920	.800	.496	.210

NOTE.—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 a value of 1 if the firm is classified as high wage during the period 1992–2000 (the years over which the Abowd-Kramarz-Margolis [AKM] estimates are obtained). Dependent variables are in logs. In col. 1, the dependent variable is individual wage. In cols. 2–5, the different firm-level outcomes are total factor productivity (TFP; output as the dependent variable, controlling for capital, materials, and labor inputs), value added, capital intensity (fixed assets as the dependent variable, controlling for firm size), and intangible capital intensity (intangible fixed assets—intellectual property, accumulated research and development investments, and goodwill—as the dependent variable, controlling for firm size), respectively. Output, value added, and capital variables are in thousands of 2000 euros and are measured over the period 1995–2001 (the years over which balance sheet data are available). Standard errors (in parentheses) are clustered by firm.

as a sample for my fixed effects estimation—equation (6). Details on sample restrictions and descriptive information are provided in appendix section 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% 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 therefore estimate equations such as

$$\ln O_{jslt} = \beta_0 + \beta_1 \mathrm{HWF}_j + \mu_{st} + \varpi_{lt} + controls_{jt} + e_{jslt}, \qquad (7)$$

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 TFP (output as the dependent variable, controlling for capital, materials, and labor inputs), value added, capital intensity (fixed assets as the dependent variable, controlling for firm size), and intangible capital intensity (intangible fixed assets—intellectual property, accumulated research and development investments, and goodwill—as the dependent variable, controlling for firm size). Columns 2–5 of table 1 show the results. In the Veneto manufacturing sector clear differences between HWFs and non-HWFs emerge: HWFs feature on average 8% higher TFP, 11% higher value added, 10% higher capital intensity, and 27% higher intangible capital intensity. This evidence is important for establishing the potential for knowledge transfer in the region.

[23] This finding emerges from the estimation of $w_{ijt} = X'_{it}\beta + \theta_i + \beta_1 \mathrm{HWF}_j + \varepsilon_{ijt}$, where the dummy HWF takes the value of 1 if firm j is classified as high wage.

B. Labor Mobility and Characteristics of Knowledgeable Workers

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, I observe around 62,000 incidents of job change. These moves are categorized according to the direction of the flows in table A3 (tables A1–A12 are available online). Column 1 shows that around 7,700 of these moves are from HWFs to non-HWFs. Columns 2 and 3 distinguish between moves within and moves across LLMs. They show that moves within the same LLM happen more frequently.[24] Columns 4 and 5 distinguish between moves within and moves between two-digit industries. Around 35% of the moves from HWFs to non-HWFs are within the same industry. The remaining moves are to a non-HWF in one of the 19 two-digit industries other than the one in which the worker has HWF experience.

Table A4 shows the share of knowledgeable workers in non-HWFs (col. 1) and the share of non-HWFs employing knowledgeable workers, that is, firms with $H > 0$ (col. 2). The proportion of knowledgeable workers in non-HWFs is defined as the number of knowledgeable workers observed at non-HWFs divided by the total number of workers in the Veneto region employed by non-HWFs (i.e., employees at non-HWFs with $H = 0$ or $H > 0$). In 1995, only 0.5% of the total 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 knowledgeable workers spread across the sample of non-HWFs. The share of firms employing knowledgeable workers is much greater than the share of such workers: around 18% in 1995 and around 29% in 2001.

Overall, I observe 6,539 unique knowledgeable workers. With regard to individual characteristics of the movers in my sample, table A5 shows that knowledgeable workers observed at non-HWF tend to be more likely to be male, white collar, and managers than non-HWF workers without recent experience at good firms.[25] For a comparison of the distribution of the estimated θ_i, see Section V.III.

V. 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 section A.II.

[24] In app. sec. 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.

A. Estimates Using Ordinary Least Squares (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 variable 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 the value of 0 for a large number of observations (fig. 1). Thus, any possible transformation of the H measure could possibly affect the associated estimated parameters.[26] Column 1 reports estimates from the baseline OLS specification: the coefficient on H_{jst} is positive (0.03) and significant. In columns 2 and 3, I use the Olley and Pakes (1996) and Levinsohn and Petrin (2003) control function methods in order to address potential endogeneity arising from unobservable productivity shocks. In implementing these control function approaches, I treat H_{jst} as a freely variable input (recall the discussion in Sec. III.B.1).[27] Although the point estimates of the coefficients for H_{jst} in the Olley and Pakes (1996) and Levinsohn and Petrin (2003) specifications 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 columns 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 (as discussed in Sec. III.B.1). In Section V.B I show results from the IV strategy. Columns 4 and 5 of table 2 show the estimates from an alternative approach 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

[26] Results using different functional forms are discussed in Sec. V.C.

[27] 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. I do not observe investment, and hence for col. 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 (i) many firm-year observations are lost when I take the difference in reported book values and (ii) the Olley and Pakes (1996) approach requires positive values for the proxy variable, eliminating additional firm-year observations (the estimation routine will truncate firms' nonpositive proxy variable observations because the monotonicity condition necessary to invert the investment function—and hence back out productivity—does not hold for these observations).

Table 2
Knowledgeable Workers and Productivity in Non-High-Wage Firms (Non-HWFs), 1995–2001

	OLS (1)	OP (2)	LP (3)	Investment-Capital Interactions in t, $t + 1$ (4)	Materials-Capital Interactions in t, $t + 1$ (5)
log(capital)	.092	.087	.148		
	(.005)	(.019)	(.010)		
log(materials)	.583	.587		.617	
	(.007)	(.007)		(.012)	
log(employees)	.223	.225	.202	.187	.177
	(.006)	(.010)	(.006)	(.014)	(.006)
H workers	.030	.018	.021	.018	.022
	(.003)	(.004)	(.003)	(.006)	(.003)
	[.003]			[.006]	[.003]
Observations	17,158	6,635	17,158	2,963	13,540
Adjusted R^2	.931			.940	.952

NOTE.—The dependent variable is log(output). "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; OP). Column 3 implements the procedure in Levinsohn and Petrin (2003; LP). 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. Standard errors in parentheses are clustered by firm, and those in brackets (for the main variable of interest) are clustered by local labor market (except in cols. 2 and 3, because the opreg and levpet Stata routines do not allow it). OLS = ordinary least squares.

that non-HWFs benefit from knowledgeable workers by experiencing increased productivity.[28]

Overall, the main empirical result discussed so far 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 between 1.8% and 3%. In interpreting my estimates, it is important to highlight that non-HWFs are quite small: the median number of employees at non-HWFs is 33. Furthermore, as many as 78% of non-HWFs in a given year do not employ any knowledgeable workers. Hiring one knowledgeable worker therefore implies a significant change in terms of workforce for most firms in my data. It may also be instructive to evaluate the average magnitude of TFP change in monetary terms. Multiplying the estimated percentage change by the mean value of non-HWF output suggests that the increase in TFP due to hiring a worker from HWFs leads to an increase in total output of €154,000–€256,000 (in 2000 euros).[29]

[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 $t + 1$).

[29] All monetary amounts in the paper are expressed in terms of the price level of the year 2000. Stoyanov and Zubanov (2012) find that the productivity gains associated with hiring from more productive firms are equivalent to 0.35% per year for

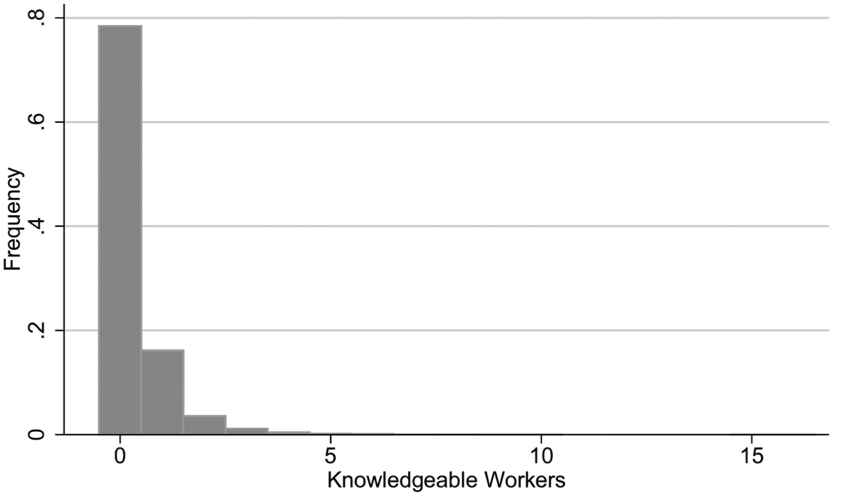

FIG. 1.—Distribution of H (number of knowledgeable workers) across firms. Shown are the number of knowledgeable workers observed at non-high-wage firms. Firm-year observations are for the period 1995–2001. I define a worker as having recent high-wage firm (HWF) experience in year t if he or she is observed 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. A color version of this figure is available online.

B. IV Estimates

In this section, I instrument for the number of knowledgeable workers in a nongood firm by using the number of good firms in the same LLM and same two-digit industry that downsized in the previous period. This IV strategy aims to address the concern of strategic mobility whereby workers may be attracted to join firms that are on the rise. Turning to the details of the instrument, a firm is defined as a downsizing firm if its employment decreased by more than 1% compared with the previous year's level. The division of good firms into downsizing and nondownsizing firms according to this criterion is less sensible for small firms. Thus, I impose the additional condition that the decrease in employment is greater than or equal to three individuals.[30] In the presence of product demand shocks or offshoring, using the number of downsizing firms as an instrument is invalid if it cannot

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 multinational enterprise experience contribute 20% more to the productivity of their plant than workers without such experience.

[30] The IV is summarized in table A7, together with other variables constructed at the LLM level that are used in the analysis.

be excluded from the model of interest (eq. [3]). The identifying assumption of my IV strategy is therefore that the number of downsizing good firms is correlated with the variable of interest, H, but uncorrelated with any other unobserved determinants of productivity. As in the above (and in the remainder of the paper), regressions include industry-year interaction dummies and LLM-year interaction dummies. The term $\widehat{\beta_H}^{\mathrm{IV}}$ is biased upward if there are localized unobservable industry shocks that would lead good firms to downsize and at the same time positively affect productivity at non-HWFs. But the IV estimates may also be downward biased because of correlated shocks (at the LLM-by-industry level) that lead some local good firms to downsize and at the same time lead to productivity declines in non-HWFs. I discuss in detail my investigation of possible violations of the exclusion restriction below.

Table 3 shows the results from the IV estimation of equation (3).[31]

The estimated coefficient of H, the number of knowledgeable workers, in column 1 is large (0.143). Recall the OLS estimates: the coefficient on knowledgeable workers is 0.03. A possible explanation for the magnitude of the IV results is that in normal times workers who leave good firms may be more peripheral workers. However, when a firm is downsizing even its core or key workers may have to leave, and these workers may have to go to worse firms (whereas if they would leave in normal times, they might go to other HWFs). Workers who move to non-HWFs due to downsizing of a HWF may therefore embody more knowledge than workers who shift in general.[32] Another, more 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. Two-stage least squares (2SLS), on the other hand, estimates the average effect for the 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.[33] If the effect of knowledgeable workers on productivity is larger for non-HWFs that are marginal in the recruitment decision, the 2SLS estimates will exceed those of consistent OLS.[34] In practice, how-

[31] Standard errors are clustered at the level of the LLM.

[32] As for observable worker characteristics, in results untabulated here I have explored whether knowledgeable workers from downsizing firms come from more high-skilled occupations. I failed to find any significant differences between the two groups.

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

[34] In results not tabulated here, I have explored whether non-HWFs hiring knowledgeable workers from downsizing HWFs differ in terms of observable characteristics from non-HWFs hiring knowledgeable workers in general. Specifically I have performed t-tests of equality of means for the levels of capital per worker,

Table 3
Knowledgeable Workers and Productivity in Non-High-Wage Firms (Non-HWFs), Instrumental Variable (IV) Estimates, 1995–2001

	Baseline (1)	Tradability (2)	Larger Drop in L (3)
H workers	.143	.143	.172
	(.067)	(.067)	(.083)
log(capital)	.085	.085	.083
	(.008)	(.008)	(.008)
log(materials)	.575	.575	.573
	(.013)	(.013)	(.014)
log(employees)	.204	.204	.199
	(.010)	(.010)	(.011)
Observations	17,158	17,158	17,158
Adjusted R^2	.914	.914	.909
Angrist-Pischke F-stat, first stage	13.82	13.81	10.39

NOTE.—The dependent variable is log(output). Standard errors (in parentheses) are clustered by local labor market (LLM; 47). Regressions include industry-year interaction dummies and LLM-year interaction dummies. Column 1 reports instrumental variable estimates using the lagged number of downsizing local good firms in the same five-digit industry. A good firm is considered to be downsizing if the drop in L is larger than 1%. 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 the value of 1 if the four-digit industry produces goods that are not widely traded outside the LLM. In col. 3, a good firm is considered to be downsizing if the drop in L is larger than 3%. The controls are the same as in col. 2. Angrist-Pischke F-stat = Angrist-Pischke multivariate test of excluded instruments F-statistic.

ever, the IV standard errors are quite large (0.067) and prevent me from drawing definitive conclusions.[35]

It is also important to emphasize that downsizing of a good firm is not a likely event. In Section VI I evaluate the extent to which labor mobility can explain the productivity advantages of firms located near other highly productive firms. Those calculations take into account the probability of downsizing and therefore deliver a similar conclusion when using either the OLS or the IV estimate of β_H to study the extent to which worker flows explain the productivity gains experienced by incumbent 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

TFP, and number of employees in the year before the hire. I failed to find any significant differences between the two groups.

[35] 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 a scenario the OLS coefficient will be biased downward because of this negative selection. Figure A6 (figs. A1–A6 are available online) does not lend support to this hypothesis, however. (The figure is for 1995, but figures for other years do not lend such support either.)

of output.[36] Unobserved shifts in local demand from the products of HWFs to the products of non-HWFs might simultaneously lead to a downsizing by HWFs, higher output prices for non-HWFs, and the 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, it is possible that $\widehat{\beta_H}^{\mathrm{IV}} > 0$ reflects higher output prices rather than higher productivity due to labor mobility. I do not expect this to be a major factor in my study: manufacturing firms in my sample generally produce goods traded outside the LLM.[37] To explore this possibility further, I add a dummy in column 2 taking the value of 1 if the industry produces goods that are not widely traded outside the LLM.[38] The results are largely unchanged.[39]

Finally, in column 3 I use an alternative definition of downsizing firms: a downsizing firm must see an employment reduction larger than 3% compared with the previous year's level. The results are again largely unchanged. Overall, the estimates suggest 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.

C. Validity and Robustness

The main empirical result so far for the firm-level analysis is that non-HWFs appear to 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.

1. Unobserved Worker Quality

As mentioned above, a potential threat to identification is the fact that I do not observe worker quality. To investigate this issue, figure 2 shows a plot of the quantiles of the distribution of $\widehat{\theta}_i's$, the worker fixed effects ob-

[36] The theoretically correct dependent variable in a productivity study is the quantity of output, but due to data limitations this study (and most of the empirical literature on productivity in a large sample of firms) uses price multiplied by quantity.

[37] 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.

[38] See app. sec. A.V for details.

[39] In results untabulated here, I have included an interaction between the dummy for local industry and H: the coefficient on the interaction is not significant (coefficient: 0.376; standard error: 0.428). I have also run separate IV regressions in the groups of tradable and nontradable industries: while in the first subsample the estimates are very similar to the baseline IV specification on the full sample (col. 1 of table 3), in the second, much smaller subsample (706 observations) $\widehat{\beta}_H$ is positive but smaller and not significant (coefficient: 0.057; standard error: 0.215).

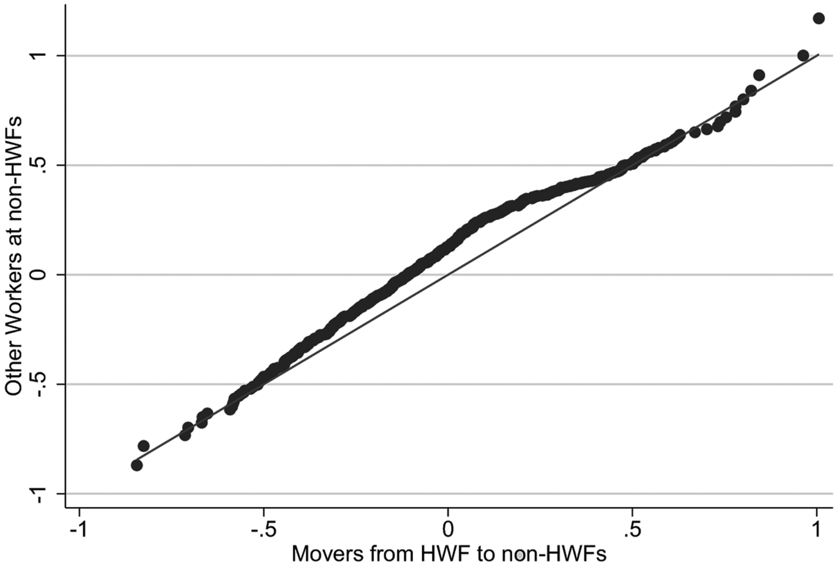

FIG. 2.—Quantile-quantile plot: worker effects. Shown is a quantile-quantile plot of worker effects in 1995. Plots for all other years are very similar. HWF = high-wage firm; non-HWF = non-high-wage firm. A color version of this figure is available online.

tained from estimating equation (6), for unknowledgeable workers (workers at HWFs with no experience at good firms) against the quantiles of the distribution of $\widehat{\theta}_i$'s's for the switchers from good firms. Points on the right-hand side of the 45° line mean that the values of the distribution on the X-axis are higher than those of the distribution on the Y-axis.[40] Since many points are on the left-hand side of the main diagonal, it seems reasonable to conclude that workers coming to non-HWFs from HWFs are (to a large extent) not positively selected on unobserved ability. Below I revisit this question in a regression framework and again conclude that the productivity effect of knowledgeable workers does not appear to be driven by unobserved worker quality.

2. *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

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

productivity gains may be driven by better worker-firm matching rather than knowledge transfer. 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 by N: the number of new hires that do not have HWF experience. I then estimate for the sample of MWFs

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

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 4 shows the results. The coefficient on H is positive and significant. The coefficient on 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. The introduction of N, in the spirit of Balsvik (2011), can also be seen as a placebo test at the firm level. It suggests that the productivity effect attributed to knowledgeable workers is not associated with recently hired workers in general: large productivity gains from hiring seem to be realized only when new hires come from HWFs. While I cannot completely rule out 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.[41]

3. *Within Estimates and an Event Study Analysis*

Next, in column 2 of table 4 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 func-

[41] In results untabulated here, I have checked how workers from LWFs are selected compared with other workers at MWFs. Looking at quantile-quantile plots of worker effects, there is no evidence of systematic selection. In some years t-tests of equality of means show evidence that workers from LWFs come from higher-skilled occupations.

Table 4
Knowledgeable Workers and Productivity in Non-High-Wage Firms (Non-HWFs): Additional Specifications Addressing Endogeneity Concerns

	Experience HWFs/LWFs (1)	Within (2)	Workforce Characteristics (3)	Materials-Capital-Labor Interactions in t, $t+1$ (4)
log(capital)	.097	.065	.091	
	(.007)	(.005)	(.005)	
log(materials)	.585	.596	.573	
	(.011)	(.013)	(.007)	
log(employees)	.224	.060	.229	
	(.009)	(.004)	(.006)	
H workers	.022	.010	.029	.012
	(.004)	(.002)	(.003)	(.003)
Recent LWF experience	.003			
	(.002)			
$\beta_{\tilde{H}}^{\mathrm{HWF}} = \beta_{\tilde{N}}^{\mathrm{LWF}}$, p-value	.000			
Observations	8,791	17,158	17,158	13,540
Adjusted R^2	.938	.986	.933	.961

NOTE.—The dependent variable is log(output). Standard errors (in parentheses) are clustered by firm. Regressions include industry-year interaction dummies and local labor market–year interaction dummies. "H workers" is the number of knowledgeable workers currently observed at non-HWFs. Column 1 is estimated on the sample of medium-wage firms and includes workers with recent experience at high-wage firms (HWFs) and low-wage firms (LWFs). "$\beta_{\tilde{H}}^{\mathrm{HWF}} = \beta_{\tilde{N}}^{\mathrm{LWF}}$, p-value" is the p-value of the equality of coefficients of the variable "Recent HWF experience" and the variable "Recent LWF experience." Column 2 reports within estimates. Column 3 adds the shares of managers, white-collar workers, blue-collar workers, female workers, 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 (2015) approach.

tions (Eberhardt and Helmers 2010).[42] 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 the parameters. If this concern is set aside, the results show a positive and significant coefficient on H.

A potential issue in evaluating the relationship between the number of knowledgeable workers and TFP arises from the possibility of differential pretrends. Moreover there may be some interesting posthiring dynamics that are not captured by the above estimation procedures. I therefore turn to an event-study research design as in Kline (2011) and Autor (2003), whose exposition I follow here.

Specifically, the regression equation is

$$\begin{aligned}\ln(Y_{jslt}) = {}& \beta_0 + \beta_L \ln(L_{jslt}) + \beta_K \ln(K_{jslt}) + \beta_M \ln(M_{jslt}) \\ & + \sum_{\tau} \beta_\tau D_{jt}^{\tau} + \mu_{st} + \varpi_{lt} + \lambda_j + u_{jslt},\end{aligned} \tag{8}$$

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

where the D_{jt}^{τ} are a sequence of event-time dummies that equal 1 when the hire is τ years away (τ may be negative). Formally,

$$D_{jt}^{\tau} \equiv I[t - e_j = \tau],$$

where $I[.]$ is an indicator function for the expression in brackets being true and e_j is the year of the hire. Therefore, the β_{τ} coefficients characterize the time path of TFP relative to the date of the event for hiring non-HWFs, conditional on the unobserved variance components μ_{st}, ϖ_{lt}, λ_j, and u_{jslt}. If the hire is not, on average, preceded by trends in firm-specific TFP, the following restriction should hold:

$$\beta_{\tau} = 0 \forall \tau < 0.$$

The results are obtained by estimating equation (8) by OLS and adding a set of event-time dummies prior to and after the hire, together with firm, industry-year, and LLM-year fixed effects. Specifically, I add dummies for 3 and 2 years before the hire, years 0–2 after the hire, and year 3 forward (I normalize the first lead—i.e., −1 in event time—to zero). Of these six dummies, note that the first five are equal to 1 only in the relevant period, while the final dummy variable for year 3 forward equals 1 in each period beginning with the third year after the event.

Column 1 of table A8 reports the estimated β_{τ} coefficients, comparing changes in the TFP of firms that hire a knowledgeable worker with both the TFP of firms that have not yet hired a knowledgeable worker and the TFP of firms that will never hire a knowledgeable worker during my sample period. Figure 3 plots the coefficients. There is no clear pretreatment trend in the coefficients but an upward break in the TFP of non-HWFs after the hire of a knowledgeable worker.[43] Estimates with linear firm-specific time trends (displayed in col. 2) are qualitatively similar to the baseline ones.

4. *Additional Specifications Addressing Endogeneity Concerns*

Considering the differences in observable characteristics documented in Section IV.B between movers from HWFs and other workers at non-HWFs, in column 3 of table 4 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 the number of employees in both t and $t + 1$. This specification is in the spirit of

[43] In results untabulated here, I have confirmed that the estimates are similar when estimating the econometric model in col. 1 without the non-HWFs that never hire, thus exploiting only the differential timing of arrival of a knowledgeable worker among the non-HWFs that hire.

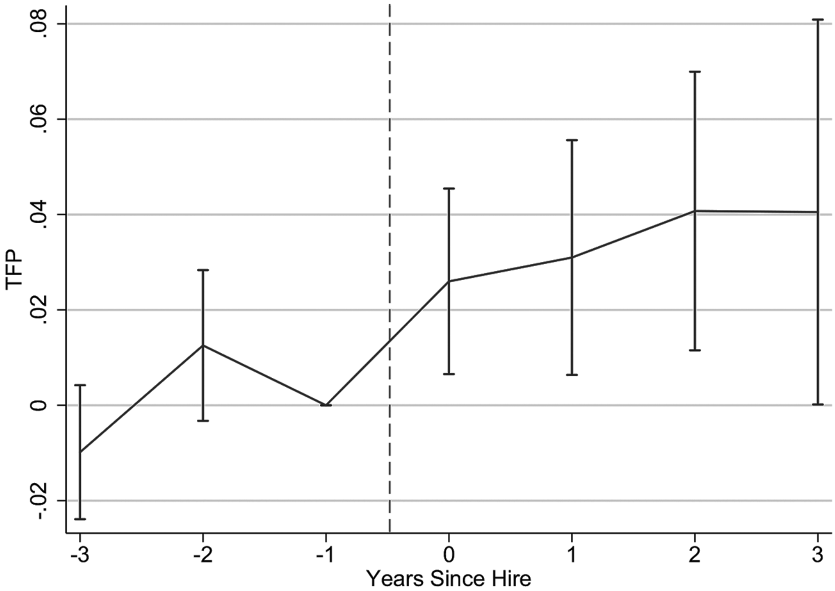

FIG. 3.—Productivity of non-high-wage firms relative to the year of the hire of a knowledgeable worker. The figure plots point estimates for leading and lagging indicators for the hire of a knowledgeable worker. Event time indicator "+3" set to 1 for all periods 3 periods after the event and 0 otherwise. The omitted category is 1 period prior to the event. Vertical bars correspond to 95% confidence intervals with firm-clustered standard errors. TFP = total factor productivity. A color version of this figure is available online.

the Ackerberg, Caves, and Frazer (2015) approach. The estimates in column 4 of table 4, together with the IV results and the estimates in columns 4 and 5 of table 2, indicate 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. *Alternative Groupings of Firms*

As an additional sensitivity check, I classify potential good firms as firms with high TFP. Specifically, I estimate firm effects from a TFP 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 A9, are very similar to those in table 2. I also experimented with a grouping of firms based on the estimated firm fixed effects in a value-added specification. Results were largely unchanged.

6. *Additional Robustness Checks*

Table A10 shows results from further robustness checks. In column 1, I show that estimates when using value added as an alternative dependent variable are qualitatively similar. 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.[44] In column 3, I estimate

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

Compared with 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 $1(H = 0)_{jslt}$, which takes a value of 1 if the number of knowledgeable workers is equal to zero.

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 2–5.

7. *Quality of the Sending Firm and Quality of the Worker*

In table 5 I extend the specification in column 1 of table 4 by estimating separate coefficients for (i) high-quality knowledgeable workers ($\beta_{\tilde{H}}^{\mathrm{HWF,h}}$), (ii) low-quality knowledgeable workers ($\beta_{\tilde{H}}^{\mathrm{HWF,l}}$), (iii) high-quality hires that do not have HWF experience ($\beta_{\tilde{N}}^{\mathrm{LWF,h}}$), and (iv) low-quality hires that do not have HWF experience ($\beta_{\tilde{N}}^{\mathrm{LWF,l}}$). In column 1 quality is measured in terms of past occupation. The difference between $\beta_{\tilde{H}}^{\mathrm{HWF,h}}$ and $\beta_{\tilde{N}}^{\mathrm{LWF,h}}$ is significant at the 1% level; that between $\beta_{\tilde{H}}^{\mathrm{HWF,l}}$ and $\beta_{\tilde{N}}^{\mathrm{LWF,l}}$ is significant at 10%. In column 2 quality is measured in terms of preestimated worker fixed effect. The difference between $\beta_{\tilde{H}}^{\mathrm{HWF,h}}$ and $\beta_{\tilde{N}}^{\mathrm{LWF,h}}$ is again significant at the 1% level; that between $\beta_{\tilde{H}}^{\mathrm{HWF,l}}$ and $\beta_{\tilde{N}}^{\mathrm{LWF,l}}$, while positive, is not significant.

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

Table 5
Knowledgeable Workers with Experience at High-Wage Firms (HWFs) versus Low-Wage Firms (LWFs): Separating Higher- and Lower-Skilled Workers

	Previous Occupation (1)	Worker Fixed Effect (2)
H workers, higher skilled	.035	.030
	(.007)	(.006)
Recent LWF experience, higher skilled	.009	.006
	(.005)	(.003)
H workers, lower skilled	.012	.008
	(.005)	(.005)
Recent LWF experience, lower skilled	.001	.002
	(.003)	(.004)
H_0: $\beta_{\tilde{H}}^{\mathrm{HWF,h}} = \beta_{\tilde{N}}^{\mathrm{LWF,h}}$, p-value	.005	.001
H_0: $\beta_{\tilde{H}}^{\mathrm{HWF,l}} = \beta_{\tilde{N}}^{\mathrm{LWF,l}}$, p-value	.077	.473

NOTE.—This table reports coefficients of interest when I extend the specification in col. 1 of table 4 by estimating separate β values for (i) high-quality knowledgeable workers ($\beta_{\tilde{H}}^{\mathrm{HWF,h}}$), (ii) low-quality knowledgeable workers ($\beta_{\tilde{H}}^{\mathrm{HWF,l}}$), (iii) high-quality hires who do not have HWF experience ($\beta_{\tilde{N}}^{\mathrm{LWF,h}}$), and (iv) low-quality hires who do not have HWF experience ($\beta_{\tilde{N}}^{\mathrm{LWF,l}}$). In col. 1, quality is measured in terms of past occupation. In col. 2, quality is measured in terms of preestimated worker fixed effect.

D. Further Firm-Level Results

1. *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 6 distinguishes between workers with HWF experience moving within the same two-digit industry and workers moving between industries. The coefficient on both types 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.

2. *Results by Worker Occupation*

I now investigate whether the occupation of new hires influences the strength of the effect of knowledgeable workers on the receiving firm's productivity. I consider heterogeneity in knowledgeable workers' occupation both within their sending (good) and receiving (nongood) firms. Specifically, knowledgeable workers are grouped into higher-skilled and lower-skilled occupations. The higher-skilled occupation category includes white-collar workers and managers. The lower-skilled occupation includes blue-collar workers and apprentices.

In column 2 of table 6, 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 occupations are better able to transfer knowl-

Table 6
Knowledgeable Workers and Productivity in Non-High-Wage Firms (Non-HWFs), Further Extensions, 1995–2001

	Same/Different Industry (1)	Previous Occupation (2)	Current Occupation (3)	Continuous Measure (4)	Lag (5)
H from same industry	.035				
	(.006)				
H from different industry	.028				
	(.004)				
H current higher-skilled occupation		.042			
		(.006)			
H current lower-skilled occupation		.025			
		(.004)			
H previous higher-skilled occupation			.044		
			(.006)		
H previous lower-skilled occupation			.025		
			(.004)		
Exposure				.026	
				(.013)	
Lag (H workers)					.024
					(.005)
$\beta_H^{\text{same}} = \beta_H^{\text{diff}}$, p-value	.232				
$\beta_H^{\text{high}} = \beta_H^{\text{low}}$, p-value		.013	.006		
Observations	17,158	17,158	17,158	17,158	16,265
Adjusted R^2	.931	.931	.931	.931	.932

NOTE.—The dependent variable is log(output). All columns include log(capital), log(labor), and log(employees). Standard errors (in parentheses) are clustered by firm. The estimation method is ordinary least squares. Column 1 differentiates between workers moving within the same industry and between industries. "$\beta_H^{\text{same}} = \beta_H^{\text{diff}}$, p-value" is the p-value of the equality of coefficients of the variable "H from same industry" and the variable "H from different industry." In col. 2, H is disaggregated into two groups based on the occupation at the previous employer (HWF). In col. 3, it is disaggregated based on the occupation at the current employer (non-HWF). "$\beta_H^{\text{high}} = \beta_H^{\text{low}}$, p-value" 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 col. 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 the main text for details on the definition of this variable). In col. 5, the variable of interest is lagged by 1 year.

edge. 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.

3. Continuous Measure of the Receiving Firm's Exposure to New Knowledge

In column 4 of table 6, I employ a continuous measure of the receiving firm's exposure to knowledge, drawn from Stoyanov and Zubanov (2012), that exploits the differences between sending and receiving firms. This procedure extends the analysis above, which uses a dummy indicating experience at a HWF. The new measure of a firm's exposure to knowledge is calculated for each firm j hiring at time t as follows:

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

where $\hat{\psi}_i^s$ and $\hat{\psi}_j^r$ 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 1 if $(\hat{\psi}_i^s - \hat{\psi}_j^r) > 0$ and 0 otherwise; the new measure is the difference in quality between the sending and the 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.[45]

4. *Lagged Number of Knowledgeable Workers*

In column 5 of table 6, I lag the number of workers with HWF experience. The coefficient is again positive and significant.

5. *Additional Firm-Level Analysis*

Highly knowledgeable workers may transfer knowledge or know-how not directly reflected in TFP. For example, if highly productive firms have different management practices as reflected in their input intensity choices, knowledgeable workers may transfer these as well. Table 1 shows that HWFs have greater capital and intangible capital intensity. Is there a relationship between the hiring of knowledgeable workers and the intensity of these inputs at non-HWFs? To investigate this issue I have estimated the model from table 2 using capital per worker and intangible capital per worker as the dependent variable (in logs) instead of the value of production (and excluding the inputs as covariates). The coefficient on H_{jst} is positive in both specifications (0.024 and 0.043, respectively) and highly significant.

In results not reported here, I have also explored, using interaction terms, whether the coefficient on knowledgeable workers is larger if the receiving firm has (i) high intangible capital intensity, (ii) high capital intensity, and (iii) large firm size. I failed to find any significant differences. I have also investigated further heterogeneity at the sending firm level by distinguishing between workers from HWFs with above versus below median (i) intangi-

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

ble capital intensity, (ii) capital intensity, and (iii) firm size. Once again, I failed to find any significant differences.

VI. Worker Flows and Agglomeration Advantages

In this section, I evaluate the extent to which labor mobility can explain the productivity advantages of firms located near other highly productive firms. In an influential study, Greenstone, Hornbeck, and Moretti (2010) find that, following the opening of a large manufacturing plant, the TFP of incumbent plants in the US counties that were able to attract these large plants increased significantly relative to the TFP of incumbent plants in counties that survived a long selection process but narrowly lost the competition. Data limitations prevented Greenstone, Hornbeck, and Moretti (2010) from drawing definitive conclusions regarding the underlying mechanism. I am able to evaluate the extent to which worker flows explain the productivity advantages of agglomeration by predicting the change in local productivity following an event analogous to that studied by Greenstone, Hornbeck, and Moretti (2010) within the worker mobility framework described above. More specifically, I study the effect of 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 Greenstone, Hornbeck, and Moretti (2010).[46] The underlying idea is that employees at these good firms acquire the relevant firm's internal knowledge. After a certain amount of time (in the empirical analysis I consider a 5-year horizon) some of the knowledgeable workers switch jobs. Therefore, other local firms can gain access to the superior knowledge by hiring such workers.

I focus on a change in the number of good local firms belonging to the same industry as firm j. This is motivated by Henderson (2003), Cingano and Schivardi (2004), Moretti (2004), and Greenstone, Hornbeck, and Moretti (2010), who found that local spillovers are increasing in economic proximity. Denote the number of knowledgeable workers moving within

[46] The large plants in Greenstone, Hornbeck, and Moretti (2010) generated bidding from local governments almost certainly because there was a belief of important positive effects on the local economy. Greenstone, Hornbeck, and Moretti (2010) observe that the mean increase in TFP after the opening is larger if incumbent plants are economically close to the large plant—economic links are measured using dummy variables indicating belonging to the same industry, indicators of technological linkages, and measures of worker flows at the industry level. Furthermore, the TFP effect increases over time. These 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 Greenstone, Hornbeck, and Moretti (2010) as "good" plants, and in order to simulate their experiment I consider a change in the number of good firms such that the change in output in a given Veneto LLM is comparable.

industry observed at firm j by H^{ind}. An overview of my procedure (in three steps) is as follows.

1. I estimate the effect on H^{ind} of a change in the number of good local firms belonging to the same industry as firm j.
2. Using the estimates from step 1, I predict the change in H^{ind} that each of the non-HWFs in a LLM would experience if a positive agglomeration shock within the local industry similar to the one considered by Greenstone, Hornbeck, and Moretti (2010) were to occur. I then multiply the predicted change in H^{ind} by $\widehat{\beta_H}^{\text{ind}}$, the estimated coefficient on H^{ind} in the firm-level 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 Greenstone, Hornbeck, and Moretti (2010).
3. I compare my estimate of the predicted contribution of worker flows to productivity changes with Greenstone, Hornbeck, and Moretti (2010)'s estimate of the overall productivity effect. This comparison allows us to get 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 LLM.

A. Step 1: Dynamic Effect on the Number of Knowledgeable Workers of a Change in the Number of Good Firms

I now discuss the issues related to the implementation of the first step, that is, 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. In practice, as I explain in detail below (under step 2), I consider the predicted change in H that a typical non-HWF would experience 5 years after a large change in the number of good local firms. 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 the count of knowledgeable workers from year $t-g$ until t. This implies that H^{ind} may exhibit some degree of persistence and suggests estimation of a dynamic model.

Consider a model of the form

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

$$e_{jlst} = m_j + v_{jlst},$$

$$E\left[m_j\right] = E\left[v_{jlst}\right] = E\left[m_j v_{jlst}\right] = 0, \tag{10}$$

where $Good\_Firms_{ls(j)t}$ is the number of local good firms at time t in the same industry as firm j. Recall that the superscript "ind" represents 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 $H^{\text{ind}}_{jsl,t-1}$ 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):

$$\Delta H^{\text{ind}}_{jlst} = a\Delta H^{\text{ind}}_{jsl,t-1} + b\Delta Good\_Firms_{ls(j)t} + \Delta v_{jlst}. \qquad (11)$$

The firm effects have now disappeared, but the lagged dependent variable is still potentially endogenous as the $H^{\text{ind}}_{jsl,t-1}$ in $\Delta H^{\text{ind}}_{jsl,t-1} = H^{\text{ind}}_{jsl,t-1} - H^{\text{ind}}_{jsl,t-2}$ 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 system generalized method of moments (GMM) estimator, which I employ below, the orthogonality conditions for the differenced equation (12) are augmented by the orthogonality conditions for the level equation (9).

In principle, another challenge in estimating equation (12) 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 several factors, including transportation infrastructure, cheap availability of local inputs, and so on. Some of these factors, whose value may be different for firms in different industries, are unobserved, and they may be correlated with $\Delta H^{\text{ind}}_{jlst}$. Therefore, $\Delta Good\_Firms_{ls(j)t}$ is treated as endogenous.

Table 7 gives the results of estimating equation (12) for the period 1992–2001.[47] 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.[48] The p-value of the Hansen test for overidentifying restrictions does not suggest misspecification. The regression shows a positive and significant coefficient on the number of good local firms ($\hat{b}$), in line

[47] Since these specifications do not require information collected from AIDA balance sheets, the sample period is not restricted to post-1995 observations. I include time dummies in order to remove universal time-related shocks from the errors.

[48] 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.

Table 7
Number of Local High-Wage Firms (HWFs) in Same Industry and Knowledgeable Workers Moving within Industry, System Generalized Method of Moments (GMM) Estimates, 1992–2001

	Baseline (1)	Two Step (2)	Lags Up to 4 (3)	Lags Up to 5 (4)
Lag (H from same industry)	.144	.136	.147	.159
	(.0719)	(.0653)	(.0717)	(.0660)
Local HWFs in same industry	.009	.009	.010	.009
	(.0012)	(.0012)	(.0015)	(.0014)
Observations	25,688	25,688	25,688	25,688
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	.304	.196	.337	.460
Hansen p-value	.321	.321	.607	.941

NOTE.—The dependent variable is "H from same industry," the number of H workers who moved into a non-HWF from a HWF belonging to the same industry. Standard errors (in parentheses) are clustered by local labor market. 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 cols. 1 and 2, for all variables only the shortest allowable lag is used as instrument. In cols. 3 and 4, lags up to 4 and 5 are used, respectively. AR(1)z, AR(2)z, and AR(3)z indicate the Arellano and Bond (1991) test of first-, second-, and third-order serial correlation, distributed as $N(0, 1)$. Hansen p-value = p-value of the Hansen test of overidentifying restrictions.

with the descriptive evidence discussed above of an important role of geographic and economic proximity in determining worker mobility. Column 1 of table 7 also shows a positive and significant coefficient for the lagged dependent variable ($\hat{a}$). The economic significance of $\hat{a}$ and $\hat{b}$ is described below (see the discussion of table 8 under step 2).

In column 2, I estimate the model with two-step system GMM and Windmeijer-corrected (Windmeijer 2005) cluster-robust errors.[49] In columns 1 and 2, for all variables only the shortest allowable lag is used as an instrument. In columns 3 and 4, I estimate the same specifications as in columns 1 and 2, including lags up to 4 and 5, respectively. The estimates in columns 2–4 are similar to those in column 1.

B. Step 2: Labor Mobility and LLM Productivity

Having estimated the dynamic effect on H_j^{ind} of a change in $Good\_Firms_{ls(j)t}$, I can predict the change in H_j^{ind}—and hence in productivity—that a non-HWF in Veneto would experience after an output increase similar to the one considered by Greenstone, Hornbeck, and Moretti (2010). As it turns out, the large manufacturing plants whose openings are studied by Greenstone, Hornbeck, and Moretti (2010) are much larger than the typical good firm

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

Table 8
Worker Flows and Agglomeration Advantages

	OLS (1)	OP (2)	LP (3)	Investment-Capital Interactions in t, $t+1$ (4)	Investment-Materials Interactions in t, $t+1$ (5)	IV (6)
$\widehat{\beta_H}^{\text{ind}}$	.036	.037	.031	.022	.026	.121
Probability of HWF downsize						.178
$\Delta\widehat{TFP}^{\text{ind,5 years}} = \widehat{\Delta H}^{\text{ind,5 years}} \cdot \widehat{\beta_H}^{\text{ind}}$	.021	.022	.018	.013	.015	.013
$\Delta\widehat{TFP}^{\text{ind,5 years}}$/overall agglomeration effect	.125	.128	.107	.076	.090	.075

NOTE.—This table provides a summation of the predicted change in productivity that is attributable to worker flows 5 years following a local output increase. The predicted changes are calculated for each of the different functional forms, i.e., $\widehat{\beta_H}^{\text{ind}}$ is obtained using ordinary least squares (OLS), the Olley and Pakes (1996; OP) approach, the Levinsohn and Petrin (2003; LP) approach, polynomial functions of both capital and investments and capital and materials, and the instrumental variable (IV) approach. Simulating results to correspond to the large plant opening results found in Greenstone, Hornbeck, and Moretti (2010) 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 Greenstone, Hornbeck, and Moretti (2010). HWF = high-wage firm; TFP = total factor productivity.

in Veneto.[50] To observe a change in local output comparable to the typical output increase caused by the opening of one large plant in Greenstone, Hornbeck, and Moretti (2010), a Veneto locality must experience an increase of 56 HWFs. This is the shock I consider for my calculations of the effect of labor mobility on LLM productivity.

The predicted change in H that a typical non-HWF would experience after 5 years, the time horizon considered in Greenstone, Hornbeck, and Moretti (2010), is then $\widehat{\Delta H}^{\text{ind,5 years}} = 56 \cdot (b + ab + a^2b + a^3b + a^4b + a^5b)$. This change in H can be obtained using the estimates for a and b in equation (12) from table 7.[51] To obtain the predicted change in productivity, I estimate a variant of equation (3) in which the variable of interest is the number of knowledgeable workers moving within industry (H^{ind}).[52] The predicted change in productivity attributable to worker flows 5 years after the local output increase is then equal to $\Delta\widehat{TFP}^{\text{ind,5 years}} = \widehat{\Delta H}^{\text{ind,5 years}} \cdot \widehat{\beta_H}^{\text{ind}}$. In the case of the IV, the number of new entrants is multiplied by the probability of

[50] This is due both to the fact that new entrants in Greenstone, Hornbeck, and Moretti (2010) are significantly larger than the average new plant in the United States and to 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 the United States. See app. sec. A.VI for descriptive statistics.

[51] I use the estimates in col. 1, which imply $\widehat{\Delta H}^{\text{ind,5 years}} = 56 \cdot 0.011 = 0.589$. In results untabulated here I have also estimated eq. (12) without the lagged term of knowledgeable workers, lagging *Good_Firms* by 5 years and without using the dynamic panel IV estimator. The coefficient on lagged *Good_Firms* is significant and equal to 0.019, yielding $\widehat{\Delta H}^{\text{ind,5 years}} = 56 \cdot 0.019 = 1.043$.

[52] The estimates of $\widehat{\beta_H}^{\text{ind}}$ using the different approaches (baseline OLS, Olley and Pakes [1996], Levinsohn and Petrin [2003], polynomial functions of capital and investments or capital and materials in t and $t + 1$, and IV) are shown in table A11.

downsizing. Table 8 provides a summation of the calculations concerning the effect of labor mobility on LLM productivity. The predicted change in productivity attributable to worker flows 5 years following a large local output increase ranges from 1.3% to 2.2% depending on the specification.

C. Step 3: Comparison with Greenstone, Hornbeck, and Moretti (2010)

The final step is to compare the magnitude of $\widehat{\Delta TFP}^{\text{ind,5 years}}$ with Greenstone, Hornbeck, and Moretti (2010)'s estimate of the overall productivity effect caused by a local output increase. The increase in productivity estimated by Greenstone, Hornbeck, and Moretti (2010) 5 years after the opening for incumbent plants in the same two-digit industry is 17%. Hence, my calculations indicate that worker flows explain 8%–13% of the agglomeration advantages estimated by Greenstone, Hornbeck, and Moretti (2010), with the mean of the point estimates being 10%.

VII. Further Analysis

A. Additional LLM-Level Analysis

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 LLM.

Another explanation that has been proposed in the literature is the possibility of advantages 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). 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 by both the number of good firms and the number of nongood firms in industry s and LLM l:

$$\begin{aligned}\ln(Y_{jslt}) = {} & \tilde{\beta}_0 + \tilde{\beta}_L \ln(L_{jslt}) + \tilde{\beta}_K \ln(K_{jslt}) \\ & + \tilde{\beta}_M \ln(M_{jslt}) + b_G Good\_Firms_{ls(j)t} \\ & + b_N Nongood\_Firms_{ls(j)t} + \varrho_{jslt}.\end{aligned} \tag{12}$$

I use both OLS and the system GMM estimator proposed by Blundell and Bond (1998). When using the latter, both the number of good firms and the number of nongood firms are treated as endogenous (I experiment with different lags of the instruments). The results are shown in table 9. The number of good firms is positively and statistically significantly related to an increase in the productivity of non-HWF j. A 1 standard deviation change in *Good_Firms* is associated with a TFP change of 1.4%–3.5% after 5 years for a given firm j. Multiplying the estimated percentage change by the mean value of non-HWF output, this calculation indicates that the increase in

Table 9
Number of Local High-Wage Firms (HWFs) and Productivity, 1995–2001

	OLS (1)	System GMM (2)	Two-Step System GMM (3)
log(capital)	.0940	.0176	.0250
	(.0050)	(.0586)	(.0695)
log(materials)	.5849	.6275	.6215
	(.0109)	(.0463)	(.0555)
log(employees)	.2295	.0034	.0044
	(.0104)	(.0109)	(.0117)
Lag 5 (local HWFs in same industry)	.0028		
	(.0007)		
Lag 5 (local non-HWFs in same industry)	−.0008		
	(.0003)		
Lag 1 log(output)		.9839	.9766
		(.0545)	(.0559)
Lag 1 log(capital)		−.0274	−.0320
		(.0527)	(.0650)
Lag 1 log(materials)		−.6140	−.6077
		(.0502)	(.0573)
Lag 1 log(employees)		.0095	.0118
		(.0154)	(.0184)
Local HWFs in same industry		.0009	.0010
		(.0006)	(.0006)
Local non-HWFs in same industry		−.0010	−.0010
		(.0006)	(.0006)
$\beta_{\text{HWFs}} = \beta_{\text{non-HWFs}}$, p-value	.000	.076	.072
Observations	17,158	13,501	13,501
AR(1)z		−11.88	−10.60
AR(2)z		.782	.827
AR(3)z		1.954	1.969
AR(4)z		−1.147	−1.169
Hansen p-value		.874	.874
Adjusted R^2	.931		

NOTE.—The dependent variable is log(output). Standard errors (in parentheses) are clustered by local labor market. Regressions include year dummies. Column 1 reports ordinary least squares (OLS) estimates. Column 2 reports system generalized method of moments (GMM) estimates. Column 3 reports two-step system GMM estimates using Windmeijer-corrected standard errors. In cols. 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, and AR(4)z indicate the Arellano and Bond (1991) test of first-, second-, third-, and fourth-order serial correlation, distributed as $N(0, 1)$. Only the shortest allowable lag is used as instrument. "$\beta_{\text{HWFs}} = \beta_{\text{non-HWFs}}$, p-value" 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." Hansen p-value = p-value of the Hansen test of overidentifying restrictions.

TFP due to a typical change in the number of local good firms is associated with an increase in total output of €198,000–€219,000 (in 2000 euros).[53] The coefficient of the number of nongood firms is either negative and significant or insignificant depending on the specification. The difference in productivity effects associated with each type of firm is significant. This introduction of *Nongood_Firms* can also be seen as a placebo test at the 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.

B. Wages of Workers at Non-HWFs

Given that the sending firms do not receive any compensation from the receiving firms, the presence of knowledge transfer through worker flows arguably implies a positive externality (Stoyanov and Zubanov 2014, 17). In table 10 I investigate how much of this externality is transferred to the employees in the form of higher wages.

First, I attempt to evaluate the wage premium for movers from HWFs to non-HWFs. Column 1 shows that the premium, conditional on worker characteristics (tenure, tenure squared, age, age squared, a dummy variable for manager and white-collar status, and interaction terms between gender and other individual characteristics), is on average 7.3%. The mean annual wage in the sample is about €19,300 (in 2000 euros), obtained after multiplying the mean daily wage (€62; see table A12) by the number of yearly working days (312). When expressed as a percentage of the mean, it appears that the wage premium is small in monetary terms with respect to the increase in TFP calculated in Section V.A. This evidence is only suggestive, but it is consistent with the finding of Stoyanov and Zubanov (2014) for Denmark, who discuss two potential explanations: "First, the hiring firms are unable to precisely identify the source of these gains due to uncertainty regarding the spillover potential of their workers. . . . Second, firms do observe the contribution but do not pay their competitive wage because of other labor market imperfections" (28).

[53] An interesting question is whether subsidies used to attract manufacturing plants are close to optimal, too large, or too small. This exercise is difficult to accomplish within the current context because of the lack of systematic information on the size of the subsidies offered by local governments in Veneto to plants to locate within their jurisdictions.

Table 10
Wages of Workers at Non-High-Wage Firms (Non-HWFs)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Experience at HWF	.073	.047	.068				
	(.003)	(.003)	(.005)				
Higher skilled × experience		.089					
		(.009)					
H workers				.012	.007		
				(.001)	(.001)		
Lag 5 (local HWFs in same industry)						.002	.002
						(.001)	(.000)
Industry × year dummies	No	No	Yes	No	Yes	No	Yes
LLM × year dummies	No	No	Yes	No	Yes	No	Yes
Firm characteristics	No	No	Yes	No	Yes	No	Yes
Observations	1,005,082	1,005,082	686,635	681,743	681,743	681,743	681,743
Adjusted R^2	.504	.491	.533	.456	.533	.452	.533

NOTE.—The dependent variable is log of individual wage. "Higher skilled × experience" is the interaction between "Experience at HWF" and the higher-skilled occupation category. The higher-skilled occupation category includes white-collar workers and managers. Worker characteristics (always included) are tenure, tenure squared, age, age squared, a dummy variable for manager and white-collar status, and interaction terms between gender and other individual characteristics. Firm characteristics are log employment, log(capital per unit of output), share of females, managers, blue-collar and white-collar workers, and differently aged workers. I exclude worker-year observations with remarkably high or low values for wages (I trim observations outside the 1%–99% range). Columns 4–7 only include workers without HWF experience. "*H* workers" is observed over the period 1995–2001, as in table 2. Standard errors are in parentheses (clustered by worker in cols. 1–3, by firm in cols. 4 and 5, and by local labor market [LLM] in cols. 6 and 7).

Column 2 show that the premium is 9 percentage points higher for knowledgeable workers in higher-skilled occupations than for those in lower-skilled occupations. Column 3 show that the results are similar when including firm characteristics, industry × year, and LLM × year dummies.

Next, I investigate the relationship between the number of knowledge workers and the wage of workers without HWF experience. The estimated coefficient for H shown in columns 4 and 5 (0.007–0.012) represents a significant fraction of the corresponding coefficients in the various TFP specifications of table 2. This finding is in line with the results in Van Reenen (1996) who, using a panel of UK companies, reports that average wages at a firm increase after a successful innovation.[54]

Finally, in column 6 and 7 I estimate a wage function for worker i employed by non-HWF firm j in industry s and LLM l augmented by the number of good firms in industry s and LLM l. A 10% increase in $Good\_Firms_{ls(j)t}$

[54] The following comparison may also be instructive. In the study discussed above, Cornelissen, Dustmann, and Schönberg (2016) find that a 1 standard deviation increase in peer ability increases wages by 0.3 percentage points. The estimates in cols. 4 and 5 suggest that a 1 standard deviation increase in H is associated with an increase in wages of 0.5%–0.9%.

(i.e., an increase of about 0.3 local good firms; see table A7) is associated with a change of 0.06% after 5 years for a given worker at firm j.[55]

VIII. Conclusions

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. I implemented several empirical strategies, including control function methods, placebo tests, and an IV approach. While none of these strategies is completely conclusive with regard to identification, together they gave evidence consistent with knowledge transfer through labor mobility. My findings imply that worker flows can explain around 10% of the productivity gains experienced by incumbent firms when new highly productive firms are added to a local labor market.

References

Abowd, John M., Robert Creecy, and Francis Kramarz. 2002. Computing person and firm effects using linked longitudinal employer-employee data. Longitudinal Employer-Household Dynamics Technical Papers, Center for Economic Studies, US Census Bureau.

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. High-wage workers and high-wage firms. *Econometrica* 67, no. 2:251–333.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. Identification properties of recent production function estimators. *Econometrica* 83, no. 6:2411–51.

Agrawal, Ajay, Iain Cockburn, and John McHale. 2006. Gone but not forgotten: Knowledge flows, labor mobility, and enduring social relationships. *Journal of Economic Geography* 6, no. 5:571–91.

Alfaro, Laura, and Maggie Xiaoyang Chen. 2014. The global agglomeration of multinational firms. *Journal of International Economics* 94, no. 2:263–76.

Arellano, Manuel, and Stephan Bond. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58, no. 2:277–97.

Autor, David H. 2003. Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of Labor Economics* 21, no. 1:1–42.

Balsvik, Ragnhild. 2011. Is labor mobility a channel for spillovers from multinationals? Evidence from Norwegian manufacturing. *Review of Economics and Statistics* 93, no. 1:285–97.

[55] In the study discussed above, De la Roca and Puga (2017) find that the medium-term elasticity of earnings (after 7 years) with respect to city size is close to 0.05.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2010. Social incentives in the workplace. *Review of Economic Studies* 77, no. 2:417–58.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen. 2013. Identifying technology spillovers and product market rivalry. *Econometrica* 81, no. 4:1347–93.

Blundell, Richard, and Stephen Bond. 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87, no. 1:115–43.

Bond, Stephen R. 2002. Dynamic panel data models: A guide to micro data methods and practice. *Portuguese Economic Journal* 1, no. 2:141–62.

Brusco, Sebastiano. 1983. The Emilian model: Productive decentralisation and social integration. *Cambridge Journal of Economics* 6, no. 2:167–84.

Card, David, Francesco Devicienti, and Agata Maida. 2014. Rent-sharing, holdup, and wages: Evidence from matched panel data. *Review of Economic Studies* 81, no. 1:84–111.

Charlot, Sylvie, and Gilles Duranton. 2004. Communication externalities in cities. *Journal of Urban Economics* 56, no. 3:581–613.

Christensen, Bent Jesper, Rasmus Lentz, Dale Mortensen, George Neumann, and Axel Werweck. 2005. Job separations and the distribution of wages. *Journal of Labor Economics* 23:31–58.

Cingano, Federico, and Fabiano Schivardi. 2004. Identifying the sources of local productivity growth. *Journal of the European Economic Association* 2, no. 4:720–44.

Combes, Pierre Philippe, and Gilles Duranton. 2006. Labour pooling, labour poaching, and spatial clustering. *Regional Science and Urban Economics* 36, no. 1:1–28.

Combes, Pierre Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sbastien Roux. 2012. The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica* 80, no. 6:2543–94.

Cooper, David P. 2001. Innovation and reciprocal externalities: Information transmission via job mobility. *Journal of Economic Behavior and Organization* 45, no. 4:403–25.

Cornelissen, Thomas, Christian Dustmann, and Uta Schönberg. 2016. Peer effects in the workplace. *American Economic Review* 107, no. 2:425–56.

Dasgupta, Kunal. 2012. Learning and knowledge diffusion in a global economy. *Journal of International Economics* 87, no. 2:323–36.

De la Roca, Jorge, and Diego Puga. 2017. Learning by working in big cities. *Review of Economic Studies* 84, no. 1:106–42.

Duranton, Gilles, and Diego Puga. 2004. Micro-foundations of urban agglomeration economies. *Handbook of Regional and Urban Economics* 4:2063–117.

Eberhardt, Markus, and Christian Helmers. 2010. Untested assumptions and data slicing: A critical review of firm-level production function esti-

mators. Economics Series Working Paper no. 513, Department of Economics, University of Oxford.
Eisensee, Thomas, and David Strömberg. 2007. News droughts, news floods, and U.S. disaster relief. *Quarterly Journal of Economics* 122, no. 2:693–728.
Ellison, Glenn, Edward L. Glaeser, and William Kerr. 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review* 3:1195–213.
Fosfuri, Andrea, Massimo Motta, and Thomas Rønde. 2001. Foreign direct investment and spillovers through workers' mobility. *Journal of International Economics* 53, no. 1:205–22.
Gathmann, Christina, Ines Helm, and Uta Schönberg. 2016. Spillover effects of mass layoffs. Technical report, University College London.
Gibbons, Robert, and Lawrence F. Katz. 1991. Layoffs and lemons. *Journal of Labor Economics* 9, no. 4:351–80.
Glaeser, Edward L., and Joshua D. Gottlieb. 2008. The economics of placemaking policies. *Brookings Papers on Economic Activity* 2008, no. 1:155–253.
———. 2009. The wealth of cities: Agglomeration economies and spatial equilibrium in the United States. *Journal of Economic Literature* 47, no. 4:983–1028.
Glass, Amy Jocelyn, and Kamel Saggi. 2002. Multinational firms and technology transfer. *Scandinavian Journal of Economics* 104, no. 4:495–513.
Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2010. Identifying agglomeration spillovers: Evidence from winners and losers of large plants openings. *Journal of Political Economy* 118:536–98.
Guryan, Jonathan, Kory Kroft, and Matthew J. Notowidigdo. 2009. Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics* 1, no. 4:34–68.
Helsley, Robert W., and William C. Strange. 1990. Matching and agglomeration economies in a system of cities. *Regional Science and Urban Economics* 20, no. 2:189–212.
Henderson, Vernon. 2003. Marshall's scale economies. *Journal of Urban Economics* 53, no. 1:1–28.
Henry, Nick, and Steven Pinch. 2000. Spatialising knowledge: Placing the knowledge community of motor sport valley. *Geoforum* 31, no. 2:191–208.
Imbens, Guido W., and Joshua D. Angrist. 1994. Identification and estimation of local average treatment effects. *Econometrica* 62, no. 4:467–76.
Irarrazabal, Alfonso, Andreas Moxnes, and Karen Helene Ulltveit-Moe. 2013. Heterogeneous firms or heterogeneous workers? Implications for exporter premiums and the gains from trade. *Review of Economics and Statistics* 95, no. 3:839–49.

Kantor, Shawn, and Alexander Whalley. 2014. Knowledge spillovers from research universities: Evidence from endowment value shocks. *Review of Economics and Statistics* 96, no. 1:171–88.

Kline, Patrick. 2010. Place based policies, heterogeneity, and agglomeration. *American Economic Review* 100, no. 2:383–87.

———. 2011. The impact of juvenile curfew laws on arrests of youth and adults. *American Law and Economics Review* 14, no. 1:44–67.

Levinsohn, James, and Amil Petrin. 2003. Estimating production functions using inputs to control for unobservables. *Review of Economic Studies* 70, no. 2:317–42.

Lucas, Robert E., Jr. 1978. On the size distribution of business firms. *Bell Journal of Economics* 9, no. 2:508–23.

Markusen, James R. 2001. Contracts, intellectual property rights, and multinational investment in developing countries. *Journal of International Economics* 53, no. 1:189–204.

Mas, Alexandre, and Enrico Moretti. 2009. Peers at work. *American Economic Review* 99:112–45.

Moretti, Enrico. 2004. Estimating the external return to higher education: Evidence from cross-sectional and longitudinal data. *Journal of Econometrics* 120:175–212.

———. 2011. Local labor markets. *Handbook of Labor Economics* 4:1237–313.

Nickell, Stephen. 1981. Biases in dynamic models with fixed effects. *Econometrica* 49:1417–26.

Olley, G. Steven, and Ariel Pakes. 1996. The dynamics of productivity in the telecommunications equipment industry. *Econometrica* 64, no. 6:1263–97.

Parrotta, Pierpaolo, and Dario Pozzoli. 2012. The effect of learning by hiring on productivity. *RAND Journal of Economics* 43, no. 1:167–85.

Petrin, Amil, Brian P. Poi, and James Levinsohn. 2004. Production function estimation in Stata using inputs to control for unobservables. *Stata Journal* 4:113–23.

Piore, Michael J. 2009. Conceptualizing the dynamics of industrial districts. In *The handbook of industrial districts*. Cheltenham: Edward Elgar.

Piore, Michael J., and Charles F. Sabel. 1984. Italian small business development: Lessons for U.S. industrial policy. In *American industry in international competition*. Ed. John Zysman and Laura Tyson. Ithaca, NY: Cornell University Press.

Poole, Jennifer P. 2013. Knowledge transfers from multinational to domestic firms: Evidence from worker mobility. *Review of Economics and Statistics* 95, no. 2:393–406.

Roodman, David. 2009. How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal* 9, no. 1:86–136.

Rosenthal, Stuart S., and William C. Strange. 2003. Geography, industrial organization, and agglomeration. *Review of Economics and Statistics* 85, no. 2:377–93.

———. 2004. Evidence on the nature and sources of agglomeration economies. *Handbook of Regional and Urban Economics* 4:2119–171.

Saxenian, AnnaLee. 1994. *Regional advantage: Culture and competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.

Stoyanov, Andrey, and Nikolay Zubanov. 2012. Productivity spillovers across firms through worker mobility. *American Economic Journal: Applied Economics* 4, no. 2:168–98.

———. 2014. The distribution of the gains from spillovers through worker mobility between workers and firms. *European Economic Review* 70:17–35.

Tattara, Giuseppe, and Marco Valentini. 2010. Turnover and excess worker reallocation: The Veneto labour market between 1982 and 1996. *Labour* 24, no. 4:474–500.

Trigilia, Carlo. 1990. Work and politics in the third Italy's industrial districts. In *Industrial districts and inter-firm co-operation in Italy*. Ed. F. Pyke, G. Becattini, and W. Sengenberger, 160–184. Geneva: International Institute for Labor Studies.

Van Reenen, John. 1996. The creation and capture of rents: Wages and innovation in a panel of UK companies. *Quarterly Journal of Economics* 111, no. 1:195–226.

Waldinger, Fabian. 2010. Quality matters: The expulsion of professors and the consequences for PhD student outcomes in Nazi Germany. *Journal of Political Economy* 118, no. 4:787–831.

———. 2012. Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany. *Review of Economic Studies* 79, no. 2:838–61.

Whitford, Josh. 2001. The decline of a model? Challenge and response in the Italian industrial districts. *Economy and Society* 30, no. 1:38–65.

Windmeijer, Frank. 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics* 126, no. 1:25–51.

Yasar, Mahmut, Rafal Raciborski, and Brian Poi. 2008. Production function estimation in Stata using the Olley and Pakes method. *Stata Journal* 8, no. 2:221.