Build or Buy?
Human Capital and Corporate Diversification

Paul Beaumont Camille Hebert Victor Lyonnet*

10th December, 2020

New version - Incomplete - Please do not circulate

Abstract

We study firms' decision to enter a new sector by using their internal resources ("build") or by acquiring an existing company ("buy"). Using a novel dataset linking French employer-employee data, M&A deals, and the breakdown of sales by business line, we show that firms' human capital determines the optimal scope of the firm and that M&As play an important role in expanding firms' human capital to enter distant sectors. A firm is more likely to buy when its human capital is distant from the sector of entry, and this relationship is more pronounced when labor market frictions make it difficult to hire key employees. Firms that enter by building realize smaller entry sales and exit more quickly, especially when their human capital is distant from the sector of entry.

Keywords: M&As, (In)organic growth, Diversification, Human capital, Labor market frictions.

JEL codes: L25, J24, J30, G34.

*Paul Beaumont is at the McGill University (paul.beaumont@mcgill.ca). Camille Hebert is at the University of Toronto (camille.hebert@rotman.utoronto.ca). Victor Lyonnet is at the Ohio State University (lyonnet.1@osu.edu). We thank Manuel Adelino, Philippe Aghion, Alice Bonaime, Fabio Braggion, Edouard Challe, Gilles Chemla, Espen Eckbo, Simon Gervais, Edith Ginglinger, John Graham, Denis Gromb, Johan Hombert, Julien Martin, Adrien Matray, Isabelle Méjean, Paige Ouimet, Gordon Phillips, Manju Puri, Amiyatosh Purnanandam, Adriano Rampini, Luc Renneboog, David Robinson, Christoph Schneider, René Stulz, Geoffrey Tate, Vish Viswanathan, Ramona Westermann, Liu Yang and Sheng-Jun Xu for valuable comments and helpful discussions. We thank also seminar participants at CREST, INSEE, HEC Paris, Université Paris-Dauphine, Tilburg University, Duke University, the Ohio State University, the 5th ECGC workshop on Governance and Control, the UT Austin Corporate Finance Forum, the 28th Mitsui Labor and Finance Symposium, NFA, SDU workshop on the Firm's Life Cycle, RCEA Growth, Innovation and Entrepreneurship workshop, UNC Private Equity symposium, AFA 2020 and SFS Cavalcade 2020. This paper received the best student paper award at the 5th ECGC workshop on Governance and Control. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'avenir" program (reference: ANR-10-EQPX-17). We are responsible for all remaining errors and omissions.

1. Introduction

Conglomerates dominate the economy. Bernard, Redding, and Schott (2010) document that diversified firms account for 91% of output despite only representing fewer than half of the total number of firms in the US. In France, while only 8% of firms are diversified, these firms account for more than 95% of output. Despite the dominance of diversified firms, a debate persists in the literature over the efficiency of diversified firms compared to standalone firms. Some papers argue that conglomerates do not efficiently allocate resources across segments due to agency problems, whereas others argue that resource allocation by conglomerate firms is consistent with value maximization.[1] One key limitation of the existing literature is that it takes diversified firms' scope as given, even though this scope is constantly evolving (Broda and Weinstein, 2010; Clara, Corhay, and Kung, 2020).

This paper takes a fresh look at the issue of conglomerates efficiency by studying how firms expand their scope, i.e. how diversified firms acquire and use scarce resources to diversify into new sectors. We ask whether firms diversify by using their internal resources ("build") or by expanding their resources through the acquisition of an existing company ("buy"). Our analysis focuses on the role of human capital as a key internal resource needed to successfully operate in a new sector and analyze how internal resources determine a firm's choice between building and buying in new sectors, which ultimately determines the firm's scope. Although our analysis applies to other assets, the availability of worker-level information and the existence of specific labor market frictions motivate our focus on human capital.

We propose a simple model based on Eckel and Neary (2010) in which a firm chooses its optimal scope of activity based on its internal human capital and the profitability of entering a new sector by building or buying. The model specifies the costs and benefits of building versus buying in a new sector. This tradeoff depends on the firm's internal human capital and its "human capital (HC) distance" to the sector of entry, a term we define to characterize how close the skills of the firm's existing employees are to the new sector. When a firm builds in a new sector, its marginal cost of production in the new sector depends on its internal human

[1]For papers arguing inefficient capital allocation within conglomerates, see Lamont (1997); Shin and Stulz (1998); Scharfstein (1998); Scharfstein and Stein (2000); Rajan, Servaes, and Zingales (2000). Lang and Stulz (1994) and Berger and Ofek (1995) document that conglomerates firms trade at a discount in the stock market relative to single-segment firms. By constrast, Stein (1997); Schoar (2002); Maksimovic and Phillips (2002); Maksimovic, Phillips, and Prabhala (2011); Giroud and Mueller (2015); Tate and Yang (2016b); Hoberg and Phillips (2018) find evidence consistent with an efficient allocation of capital and labor within conglomerates.

capital and the human capital distance to the sector of entry. When a firm buys to enter a new sector, it accesses the target's productive resources but has to incur the cost of acquisition.[2] We show that firms are only willing to pay the acquisition cost if it is offset by the acquisition of the skills of the target's employees. As a result, it is more profitable for firms to buy in sectors distant from the firm's core competence and build in closer sectors. Conversely, the more key employees there are in a firm's internal human capital pool to operate in the sector of entry, the lower the adjustment costs to the new sector are, and the more profitable it is for a firm to build rather than buy. The model provides a rich set of testable implications for how the firm's core competence, the growth opportunities in a new sector, and labor market frictions interact with HC distance in determining firm's decision to build or buy.

Our first contribution is to provide the first cross-industry statistics on firms' decisions to build or buy. We use the detailed breakdown of French firms' sales by sector, which is maintained by the French Bureau of Statistics (Insee), to identify entry into new sectors.[3] We identify diversifying entries at the business group-level and flag a firm as having entered a new sector if at least one of its (newly acquired or already existing) subsidiaries begins selling in the new sector and if none of its other subsidiaries already operates in that sector. We then distinguish between build and buy entries using M&A deals retrieved from SDC Platinum and Bureau van Dijk Zephyr databases that we match to the French administrative data.[4] The entry is a "buy" if the subsidiary that begins selling in the new sector has been acquired by the firm. By contrast, the entry is a "build" if the entry is made though an existing subsidiary. Our final dataset consists of 35,000 build or buy decisions in France from 2003 to 2014. We show that 98% of entries into a new sector come from firms that build on their internal resources. The figure is 85% when weighting by entry sales, meaning that diversification by acquisition represents 15% of the universe of corporate diversifications in France.

Our model's main prediction is that a firm's human capital distance to a new sector explains its decision to build or buy. One key challenge in testing this hypothesis is to construct an empirical measure of human capital distance at the firm level. Human capital is neither directly

[2]We also consider the firm's option to complement its internal human capital by hiring new employees, which implies a costly search for employees with suitable skills for the sector of entry. Hiring costs have been estimated to account for between one-quarter and one-half of wage payments and to increase in the specificity of skills (Abowd and Kramarz, 2003; Blatter, Muehlemann, and Schenker, 2012).

[3]The detailed breakdown of firms' sales across sectors is available in the Annual Business Survey (EAE-VAC). In our main analysis, we define sectors at the 3-digit level of the French classification of industries.

[4]We develop a webcrawler that takes as input the acquiring and target firms' names and addresses to link SDC Platinum and Zephyr deals to the French administrative data.

observable nor easily defined. We rely on French matched employer-employee data to construct a firm-level variable, human capital distance, which measures the extent to which prior to entry, a firm employs workers in key occupations for the sector of entry.[5] This measure captures how distant a firm's human capital is to the employment structure of all firms in the sector of entry.

We test our model's main prediction by comparing firms that operate in the same sector of origin and diversify in the same sector of entry in the same year (using sector of origin $\times$ sector of entry $\times$ year fixed effects). This specification neutralizes potential unobservable (and potentially time-varying) synergies between the sector of entry and that of origin. In line with our model's main prediction, we find that firms are more likely to "build" in sectors that are closer and to "buy" in more distant sectors. A one-standard-deviation decrease in human capital distance is associated with a 0.4 percentage-point increase in the likelihood of buying. This relationship is sizable, equal to 20% of the unconditional probability of buying. We check the robustness of our results by considering alternative measures of human capital distance and sales thresholds to identify entries.[6]

One concern with the interpretation of our results is that a firm's human capital distance to a sector of entry may be endogenous to its choice to build or buy. Indeed, firms may adjust their employment composition in anticipation of their decision to enter a new sector, in a way that depends on whether they build or buy. To address this issue, we rely on two pseudo instrumental variable (IV) strategies. First, similar to Boehm, Dhingra, and Morrow (2019), we regress the choice to buy rather than build on the human capital distance as of the first year the firm appears in our sample. The idea is that a firm's workforce composition is exogenous to its entry decisions in the distant future. We find that a firm's maximum lagged human capital distance is positively and significantly related to the likelihood of diversifying by acquisition. However, the point estimates on the maximum lagged human capital distance are significantly lower than our baseline measure, suggesting that firms do adjust their workforce composition during the

[5] Our HC distance is related to Lee, Mauer, and Xu's (2018) measure of Human Capital Relatedness. However, our measure differs in two notable respects. First, our measure is a firm-level variable, while theirs is a sector-level variable. We are able to construct a firm-level measure because we observe the employment composition of both the firm and sector of entry. Second, we compute the distance between the firm and the sector of entry, whereas Lee, Mauer, and Xu (2018) construct a measure of similarity. See Section 3.2.2 for a discussion.

[6] Our results are robust to measuring HC distance based on (i) the similarity between the share of the firm's wage bill going to an occupation and the share of the total wage bill of all firms in the sector of entry going to the same occupation and (b) the similarity between the share of the total number of hours worked by an occupation in a firm and the share of the total number of hours worked by an occupation in all firms in the sector of entry. Our results are also robust to imposing a threshold on the firm's ratio of sales realized in the new sector relative to total sales (no threshold, minimum 1%, minimum 5%, and minimum 10%), as well as using a WLS model by entry sales instead of an OLS model.

years prior to entry, but only partially. Second, similar to D'Acunto, Tate, and Yang (2018), we instrument the firm's human capital distance by an average sector-level human capital distance. This sector-level human capital distance uses the workforce structure of all firms in the sector of origin and is therefore exogenous to an individual firm's decision to build or buy (in the spirit of Lee, Mauer, and Xu, 2018).[7] We use this instrument both in the reduced-form estimation and in a two-stage least squares (2SLS), and we find that the relationship between human capital distance and the likelihood of diversifying by acquisition remains positive and significant.

Another concern is that our main results may suffer from omitted variable bias if unobserved variables drive both the choice to build or buy *and* the firm's workforce composition (e.g., the CEO's management style). In that case, the effect of the human capital distance would capture firm-level unobservable variables. Ideally, we would like to compare firms that are similar in every aspect except their human capital composition. To address this issue, we show that our main findings are robust to various combinations of fixed effects, including size buckets and second main sector interacted with the sectors of origin-entry fixed effects. In addition, our main results hold when adding firm fixed effects or controlling for a wide variety of firm characteristics such as size, profitability, capital intensity, cash holdings, and other firm assets.

Moreover, to show support for our main finding that human capital distance matters in firms' decision to build or buy, we check that human capital distance matters over and above the role of other measures of distance between the firm's sector of origin and that of entry. Specifically, we construct two proxies for the product market and geographic distance between the firm and the sector of entry. The first proxy is a product market distance measure that captures the similarities between the firm's sector of origin and the sector of entry (e.g., production synergies or a common customer base). The second proxy captures the physical distance between the firm's location and its market of entry. We find that both distances are positively associated with a firm's likelihood of entering a given market by acquiring a company already operating in that market. However, our point estimate of the effect of human capital distance does not change significantly when adding these distances to our regressions, suggesting that the product market and geographic distance proxies are mostly uncorrelated with our human capital distance measure. Interestingly, the magnitude of the human capital distance coefficient

[7]The sector-level human capital distance is the firm's weighted average of the sector-level human capital distance. The weights correspond to the firm's exposure to sectors prior to entry into the new sector. Note that we replace the combination of origin *times* entry × year fixed effects by origin + entry + year fixed effects to allow for sufficient variation to estimate the coefficient of the sector-level HC distance.

increases when interacted with each of the two distance measures. This finding suggests that internal resources matter more when the cost of building is likely to be higher.

We provide a range of cross-sectional tests providing support for our model's tradeoff between building and buying. First, we show that the relation between human capital distance and the mode of entry is stronger when workers with specific skills are difficult to hire on the external job market, i.e., when workers in key occupations are scarce.[8] This finding is consistent with the idea that firms must search for employees who are skilled to produce in the new sector. When hiring costs increase, buying an existing firm becomes relatively more attractive than building from scratch and hiring new workers.[9] Second, firms are significantly more likely to buy in more sectors with growth opportunities. Using new firm creation as a proxy for growth opportunities, we show that firms buy relatively more in sectors close to their core competence when they face high growth opportunities. However, this effect disappears when firms' human capital is distant from the sector of entry. Finally, consistent with our model, we find that firms perform worse when their human capital is distant from the distant of entry: Their sales are 20% lower on average and their probability of exiting quickly the sector of entry increases by 35%. Nevertheless, we find evidence that these firms adjust their workforce to reduce their human capital distance following their entry.

A potential limitation of our empirical analysis is that it focuses on firm's choice between build and buy without modeling the firm's decision to diversify in the first place (Campa and Kedia, 2002). If entry into a new sector and the type of entry are joint decisions, our OLS estimates may not show a causal effect of HC distance on the decision to build or buy. This selection problem is twofold. First, the firms in our sample are selected because they choose to diversify. Our model shows that only firms that are profitable enough (i.e., efficient enough) and firms with human capital that is not overly specialized diversify. By restricting our sample to the universe of firms that diversify at least once, our results should therefore be interpreted as valid for firms that are profitable enough and/or have human capital that is not overly specialized. Second, conditionally on diversifying, firms are selected because they decide *which*

[8] We use data on worker scarcity across local labor markets at the occupation level. In the data, workers are defined as "scarce" in an occupation under two conditions: (i) job offers for this occupation exceed job applications, and (ii) surveyed employers anticipate that they will not fill in a job in this occupation.

[9] Cisco's acquisition strategy illustrates this tradeoff. In 1997, a Cisco analyst described the strategy of the firm as follows: "in today's economy, building work teams from scratch can be yesterday's luxury. So, when you can't build fast enough, you buy". By 2017, Cisco had undertaken more than 200 M&As to "provide a capability, acceleration potential or earlier sectoral entry compared to partnering or developing in-house" (Wysocki, 1997; Romanski, 2017).

sector to enter. Our model allows us to treat this selection problem by modeling a firm's choice of which new sector to enter based on its human capital distance to the sector of entry. We take our model's prediction to the data by regressing a dummy equal to one if a firm enters a given sector, zero otherwise, on our HC distance. Similar to our main specification, we use sector of origin × sector of entry × year interacted fixed effects and the same set of control variables. Additionally, we use firm × year fixed effects to capture the firm's likelihood of diversifying in a given year, allowing us to identify the direction of changes in the firm's industry mix rather than changes in the number of sectors in which the firm chooses to operate. We find support for our model's prediction that firms diversify in sectors for which their human capital is close.

Related literature. Our paper is primarily related to the literature on resource allocation within conglomerates firms. Several papers point to the conglomerate discount (among others Lamont, 1997; Scharfstein, 1998; Rajan, Servaes, and Zingales, 2000), whereas others show evidence of an efficient allocation of capital within firms (Schoar, 2002; Maksimovic and Phillips, 2002; Giroud and Mueller, 2015) or or labor (Tate and Yang, 2016b; Cestone et al., 2016), arguably due to talented managers (Maksimovic and Phillips, 2013). Consistent with this "resource-based" view of the firm (Penrose, 1955; Chandler, 1992), our paper offers a complementary explanation based on firms' human capital resources that is microfounded and testable. We show that firms are more likely to buy when the firm does not employ the key occupations needed to produce in the new sector, especially when these key occupations are scarce in external labor markets.[10]

Thus, our paper also contributes to the rapidly growing literature on labor and corporate finance.[11] Tate and Yang (2016b) predict diversifying M&As by cross-industry labor flows under the condition that human capital is transferable across sectors. Ouimet and Zarutskie (2016) and Chen, Gao, and Ma (2018) show that acquiring labor is an important motive for corporate acquisitions. Lee, Mauer, and Xu (2018) find that firms are more likely to merge and have better post-merger outcomes when the target has similar human capital. We generalize results

[10] Also consistent with the "resource-based" view of the firm, Bernard et al. (2007, 2018); Boehm, Dhingra, and Morrow (2019) find that firms tend to produce in certain pairs of industries that require similar intermediate inputs.

[11] Other papers in labor and finance study the consequences of M&As on employment and wages (Dessaint, Golubov, and Volpin, 2017; Lagaras, 2017; Ma, Ouimet, and Simintzi, 2019). Some strands of the labor and finance literature examine the role of human capital in relationship with corporate investment (e.g., Merz and Yashiv, 2007; Xu, 2018; Bai, Fairhurst, and Serfling, 2018), capital structure (e.g., Matsa, 2010; Agrawal and Matsa, 2013; Simintzi, Vig, and Volpin, 2014; Baghai et al., 2018; Serfling, 2016) and asset prices (e.g., Donangelo, 2014; Kuehn, Simutin, and Wang, 2017).

in the literature on M&A and labor and show theoretically and empirically that firm's human capital composition determines not only the choice of a given sector for M&A entry but also whether a firm builds or buys in a more or less distant sector. Our results show that a firm with human capital distant from the sector of entry's predicts entry by acquisition rather than an organic entry.

Therefore, our paper also relates to the literature on the determinants of the decision to build or buy across sectors. Very few papers have jointly studied these alternative approaches to the formation of conglomerates.[12] McCardle and Viswanathan (1994) and Yip (1982) show that firms enter a new sector through acquisition when barriers to entry are high.[13] Phillips and Zhdanov (2013) show that small firms invest in R&D, whereas large firms buy those small firms that have successfully innovated. Moreover, Bernard, Redding, and Schott (2010) document that US multiproduct firms often vary their product mix but infrequently do so through acquisition (only 7% of cases). We contribute to this small strand of the literature by documenting that, at the scale of the French economy, 90% of corporate growth is made organically, by firms that build from scratch using their preexisting resources rather than buying an incumbent in the sector of entry.

2. Theoretical Framework: "Build" or "Buy"?

We propose a model that predicts a firm's scope of activity, that is, which sectors the firm diversifies in, and whether it builds or buys in each sector based on its internal workforce composition.

[12]The large literature on corporate diversification focuses primarily on the choice of sector in the decision to diversify (Campa and Kedia, 2002; Maksimovic and Phillips, 2013). Other papers focusing on M&As specifically identify CEO overconfidence, acquirer-target relatedness, and shareholder activism as determinants of diversified M&As (Fan and Goyal, 2006; Malmendier and Tate, 2008; Rhodes-Kropf and Robinson, 2008).

[13]Yip (1982)'s empirical study supports McCardle and Viswanathan (1994)'s theoretical analysis and shows that firms are more likely to build in a sector with low barriers to entry. In another contribution related to ours, Elsas, Flannery, and Garfinkel (2014) show that M&As are more likely to be financed externally relative to capital expenditures.

2.1. Environment

We study a firm's choice to diversify in a new sector. There exist infinite sectors, in each of which the firm faces a demand function

$$q = a - bp_i, \tag{1}$$

where i denotes the sector, a captures the total profitability of the sector, b is the price elasticity of demand, and p_i denotes the firm's product price.[14] The firm production function is $Y_i = L_i$ where the sole input, labor, is denoted by L_i, and there exists a fixed cost of production, F, to produce in each sector. Sectors are ordered from closest to most distant from the firm's core sector ($i = 0$). The firm's marginal cost of production in sector i is

$$c(i) = c_0 - m_i + \gamma_k i. \tag{2}$$

The marginal cost of production in a new sector i is determined by three factors. First, the firm's core competence affects the firm's marginal cost of production in any sector i. The first term in (2), c_0, represents the firm's core competence, that is, the marginal cost of production in the firm's core sector. c_0 can be interpreted as the firm-specific internal capabilities, such as managers' human capital or organizational capital, which affect the firm's productivity in every sector in which the firm operates.

Second, the firm can decrease its marginal cost of production in the new sector by hiring additional employees specialized in that sector. If the firm chooses to do so, its marginal cost of production decreases by m_i, as shown in the second term of (2). When considering whether to build in a new sector by hiring new employees, the firm trades off the reduction m_i in its marginal cost of production against the costs associated with hiring new employees specialized in that sector. We assume that these costs are quadratic and given by $h(i)m_i^2$, where $h(i)$ captures the difficulty of hiring new employees specialized in sector i. These hiring costs reflect scarcity in the labor market for specialized employees and matching frictions. In Section 6.1 we proxy for hiring costs using occupation-level data on labor scarcity.

Third, the marginal cost of production is higher when the firm enters a sector that is distant

[14] We assume a simple demand function for tractability and to abstract from monopolistic and strategic considerations. This allows us to focus on the role of firms' resources in their decision of whether to enter a new sector and how to enter it.

from its core sector: $\frac{\partial c(i)}{\partial i} = \gamma_k > 0$. The third term in (2), $\gamma_k i$, is the increase in the firm's marginal cost when producing in a sector i distant from its core sector. We interpret γ_k as the adaptability of the firm's workforce to entering a new sector. When γ_k is high, it is difficult for the firm's workforce to adapt to producing in a new sector. For instance, a high-tech firm hires employees with very specialized skills (e.g., developers) who could not easily produce in any sector other than in the tech industry. By contrast, a firm producing clothes hires employees with more generalist skills (e.g., machine workers) who could more easily adapt to producing in a new sector, such as the footwear industry.

The novelty of our framework is to study the firm's choice between two types of entry, which determines the endogenous value of γ_k. A firm has two options when entering a new sector. First, it can "build" on its existing resources, in which case $\gamma_k = \gamma_I$. Second, the firm can "buy" in the new sector, that is, acquire an existing company already operating in this sector, in which case $\gamma_k = \gamma_E$. When buying in the new sector, the firm accesses the target firm's workforce and lowers the increase in its marginal cost of production to $\gamma_E < \gamma_I$.[15] The decrease in marginal cost of $\gamma_I - \gamma_E$ represents the synergies between the acquiring firm's core competence, c_0, and the target's workforce. To acquire a target, the acquiring firm must pay an acquisition cost to the target's owner. We assume that a fraction $(1 - \alpha_E)$ of the merged firm's profit in the new sector must be paid to the target's owner.[16]Because of the acquisition cost, the acquiring firm only finds it profitable to buy the target if there exist sufficient synergies between the acquirer and target (i.e., if $\gamma_I - \gamma_E$ is large).

2.2. Build or buy?

We start by analyzing the firm's choice to enter each sector by building or buying, taking as given the firm's initial scope of activities, that is, the set of sectors in which the firm operates. We compare the firm's profit from entering a new sector by building versus buying, taking into account the firm's decision to hire additional employees specialized in the new sector. For a

[15]The firm would never pay acquisition costs to acquire a target that has workforce adaptability such that $\gamma_E > \gamma_I$. When acquiring a target, the firm also accesses the target's intangibles, such as its brand value, which would be costly if not impossible to reproduce ex nihilo. Although synergies between the acquiring firm's core competence and the target firm's intangibles can exist, for simplicity, the model abstracts from this dimension when considering the range of sectors in which the firm buys.

[16]In other words, we assume that the acquiring firm uses stock to purchase the target. In practice, all-stock offers represent 37% of initial bids, compared to the 37% for mixed offers and 26% for all-cash offers (Betton, Eckbo, and Thorburn, 2008).

given scope, we show in Appendix A.1 that the firm's profit from building is

$$\pi^{Build}(i) = \frac{1}{4b1 - \frac{1}{h}}(a - (c_0 + \gamma_I i))^2 - F, \tag{3}$$

and the firm's profit from buying is

$$\pi^{Buy}(i) = \frac{1}{\frac{4b}{\alpha_E} - \frac{1}{h}}(a - (c_0 + \gamma_E i))^2 - F. \tag{4}$$

We compare the firm's profit from building and buying depending on the distance between the new sectors and the firm's core sector. We show in Appendix A.1 how we obtain Main Prediction 1.

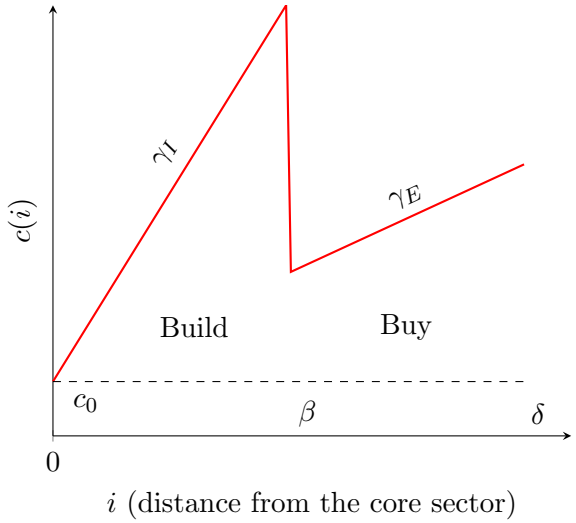
Main Prediction 1. *A firm is more likely to buy rather than build in sectors distant from its core competence.*

For a given scope, the firm buys an existing company to enter a sector distant from its core competence. Intuitively, the firm is willing to pay the acquisition cost only if it is offset by the acquisition of the specialized skills of the target's workforce. For sectors closer to the firm's core competence, building on existing resources and perhaps hiring new employees is more profitable than acquiring an existing firm and paying the acquisition premium. Denoting $\delta \geq 0$ as the most distant sector in which the firm buys and $\beta \geq 0$ as the most distant sector in which the firm builds, Figure 1 illustrates the prediction in Main Prediction 1 that a firm builds in less distant sectors and buys in more distant sectors.

Importantly, what we dub "distance" between the firm's core sector and the new sectors should be interpreted as "human capital distance." Indeed, the firm's marginal cost in (2) depends on the distance between the firm's core competence and the specialized skills in the sector of entry. We discuss in Section 3.2.1 how we empirically measure this human capital distance. In line with the model's prediction, we show in Section 5.3 that human capital distance matters over and above the role of product market and geographic distance.

Additional predictions. The model's comparative statics allow us to make several other minor predictions. First, only the most productive firms can efficiently manage their target, generating enough synergies for the share $(1 - \alpha_E)$ of the merged firm's profit to be sufficiently high for buying to be profitable. Therefore, only those firms with a high core competence (i.e.,

Figure 1. Diversification and Entry Type



firms with a production cost $c_0 \leq \underline{c_0}$), will build *and* buy to enter new sectors. Therefore, we expect firms that build and buy to be relatively more profitable and hence larger than other firms. We test this prediction in Section 3.2.1.

Second, the profit from building are more affected by the costs of hiring specialized employees (h) relative to the profit from buying. Therefore, for a given human capital distance between the firm's core sector and the new sector, a firm is more likely to enter the new sector by buying when the costs of hiring employees specialized in that sector is high (h is high). We predict the human capital distance interacts with employee hiring costs in determining whether the firm builds or buys. We test this prediction in Section 6.1.

Finally, buying implies that a fraction $(1 - \alpha_E)$ of the merged firm's profit in the new sector must be paid to the target's owner, while building does not dilute the firm's profit in the new sector. This can be seen by noting the multiple α_E in the profit from buying in (4) and the absence thereof in the profit from building in (4). For a given human capital distance between the firm's core sector and the new sector, a firm is less likely to enter the new sector by buying when the profitability of that sector is high. Thus, we predict the human capital distance interacts with growth opportunities in the new sector in determining whether the firm builds or buys. We test this prediction in Section 6.1.

2.3. Optimal firm scope

We showed that the firm will either build in close sectors and buy in distant sectors, only build in close sectors, or not diversify at all. In Appendix A.2, we show how firms' core competences determine what firms optimally choose which strategy.

Main Prediction 2. *A firm is less likely to diversify in sectors distant from its core competence.*

Only firms that are productive enough (i.e., with a high enough core competence) will diversify (i.e., firms with a production cost $c_0 \leq \overline{c_0}$, where $\overline{c_0}$ is given in Appendix A.2) because only these firms are efficient enough for the profit in the new sector to offset the fixed costs of production. The intuition for Main Prediction 2 is that the more distant the new sector is, the less likely a firm will have a sufficiently high core competence to find it profitable to enter the new sector.

3. Data and Empirical Strategy

3.1. Data sources and main variables

Our dataset consists of the merger of M&A datasets with several administrative datasets available from the French Bureau of Statistics (Insee).

3.1.1 M&A deals

We collect all M&A deals from SDC Platinum and Bureau van Dijk Zephyr between January 2003 and December 2014. We exclude leveraged buyouts and private equity deals from the sample. We focus on majority deals in which the acquirer owns less than 50% of the target shares before the acquisition and more than 50.1% thereafter. We retain deals that involve a French acquirer and a French target to be able to match them with the French administrative data.

SDC Platinum and Bureau van Dijk's Zephyr do not provide the standardized French firm identifiers. We proceed in several steps to find them.[17] These steps allow us to identify firms

[17]First, we use tickers (available only for publicly traded firms) and the Bureau van Dijk identifiers (available only for deals in Zephyr) to find a fraction of the firm identifiers. Second, we build a Python webcrawler on two websites, which takes as inputs a firm's name and address: www.bodacc.fr (Bulletin Officiel des Annonces

involved in acquisition deals at the business group level. Some of the subsidiaries may share a common name with their parent company. Our final sample includes 7,165 deals from 2003 to 2014 involving 4,139 acquiring firms.

3.1.2 Firm-level data.

Our level of analysis is the business group.[18] A business group includes the parent company and majority-owned subsidiaries connected though ownership links. We use the ownership link dataset (*Enquête sur les Liaisons financières entre sociétés*, LIFI) to retrieve the structure of groups and identify all entities linked to the diversifying business group. We also use LIFI to identify all the subsidiaries of M&A targets. We treat these subsidiaries as if they were themselves acquired along with their parent company.

The tax files (*Bénéfices Industriels et Commerciaux*) provide detailed yearly accounting information (balance sheet and income statements) at the entity level.[19] We retrieve sales, cash, total assets, tangible assets, and the firm's main sector of activity from the tax files. Sectors are defined by the French classification of sectors (*Nomenclature des activités Francaises, NAF*). We define sectors at the 3-digit industry level, which include 272 different sectors.[20] We use the SIRENE registry to obtain the geographic locations of firms' plants. We then consolidate all the different relevant variables at the group level.

Build or buy? To identify entries into new sectors, we use the subsidiaries' breakdown of sales by sector (*Enquête Annuelle des Entreprises - EAE*, then *Ventilation des Ventes par Activité* -

Civiles et Commerciales), which is a governmental website that reports official notifications since 2003, and www.societe.com, which is a commercial website that aggregates information about French companies from various sources (mostly from the French Bureau of Statistics and Bodacc.fr). Both websites are supposed to cover the universe of French firms. Third, after running the webcrawler, we drop companies for which the address, city and zip code are missing because we cannot identify with certainty the corresponding company identifier among several matches. We retain only observations for which the Jaró-Winkler string distance between the original name and the retrieved name is above 0.8. Fourth, we manually check the resulting matches. In addition, note that we cannot use changes in ownership links to identify M&As deals, as we cannot distinguish between newly reported ownership links in the database and changes in existing ownership links.

[18] French administrative data include three levels of firms: the group level, identified with an *sirtg* identifier, the subsidiary or entity level, identified with a *siren* identifier, and the establishment level, identified with a *siret* identifier.

[19] We exclude firms in the financial, agricultural and public sectors as they use different accounting systems. We focus on firms that use the standard tax forms ("Régime normal") and discard firms that use the simplified version ("Régime simplifié") because some important variables are not reported in the simplified regime. Firms in the standardized regime represent on average 94% of total value added.

[20] Using the 3-digit level of the French classification of sectors allows us to have stable categories over the sample period. Moreover, adopting a rather coarse level of sector classification limits the risk of mistaking a reporting error in the sector of activity for an entry into a new sector.

VAC). This dataset records the detailed number of sales realized by each entity in every sector every year.[21]

We flag a buy entry into a new sector, i.e., an entry by acquisition, if the firm acquires an entity in a sector in which it has not already been operating before. We also check that the entity that reports sales in the new sector ("entering subsidiary") becomes one of the acquirer's subsidiaries only after the M&A deal. In other words, we identify a firm's entry into a new sector if at least one of its subsidiaries begins selling in that sector and if none of its other subsidiaries already operate in the sector. By contrast, a firm builds or enters organically if the entering subsidiary was already controlled by the firm before the entry.

3.1.3 Matched employer-employee dataset

We use the French matched employer-employee administrative dataset (*Déclarations Annuelles des Données Sociales*, DADS) to construct the firm-level measure of human capital distance to the sector of entry. When filing payroll taxes, employers are required to report every year detailed information about their employees. Employers report the type of contract, gross and net wages, the number of hours worked and an occupation code for each worker.[22] Occupations are reported as 4-digit codes (*Nomenclatures des professions et catégories socio-professionnelles des emplois salariés des employeurs privés et publics*, PCS-ESE). The occupation codes consist of 414 different occupations, including, for instance, 28 different types of engineers (e.g., logistics, IT, electrical, or mechanical). We start the analysis in 2003, as occupation codes are often missing before this date.

We define human capital distance as follows:

$$\text{HC Distance}_{g,n,t} = 1 - \frac{\sum_i s_{g,i,t} \cdot s_{n,i,t}}{\sqrt{\sum_i s_{g,i,t}^2}\sqrt{\sum_i s_{n,i,t}^2}}, \tag{5}$$

where $s_{g,i,t}$ is the share of employees in firm g employed in occupation i and $s_{n,i,t}$ the share of employees in sector n employed in occupation i.[23] HC distance$_{g,n,t}$ ranges from 0 to 1, with 1

[21] The data are available at the entity level and exhaustive for entities with at least 20 employees. It randomly includes smaller firms. However, the survey still covers at least 85% of sales realized in a given sector. We check that subsidiaries included in the survey cover 96% of sales in the manufacturing sector. For smaller subsidiaries for which sales breakdowns are not available in the survey, we make the assumption that these subsidiaries sell only in their main sector of activity, and these data are retrieved from the tax files.

[22] Note that reporting the occupation code is required for firms with at least 20 employees and is optional for firms below this threshold.

[23] Although, we observe firms' employment compositions and breakdowns of sales by sector, one limitation of

being the maximum distance between firm g's human capital and that of sector n. In contrast, a distance of 0 indicates a perfect overlap between firm g's human capital and that of sector n.

3.1.4 Other data sources

Labor scarcity. We measure labor scarcity at the commuting zone level using data from the French national unemployment agency (*Pôle emploi*). The unemployment agency started to track scarce occupations within 350 different commuting zones in 2010. Occupations are flagged as scarce when labor demand exceeds labor supply for a given occupation, or when this occupation is anticipated to be scarce according to a survey of employers conducted by the unemployment agency.

Using the list of scarce occupations by commuting zone, we construct a firm-level measure of labor scarcity as follows:

$$\textit{Labor scarcity}_{g,n,t} = \sum_{z} \omega_{g,z,t} \cdot \textit{Zone labor scarcity}_{z,n,t} = \sum_{z} \omega_{z,g,t} \cdot \sum_{i} \textit{Scarce occupation}_{i,z,t} \cdot s_{n,i,t},$$

where $\omega_{z,g,t-1}$ is the share of sales realized by firm g in zone z in year t, $\textit{Scarce occupation}_{i,z,t}$ is a dummy variable that identifies a scarce occupation i in commuting zone z in year t, and $s_{n,i,t}$ the share of employees in sector n employed in occupation i. $\textit{Labor scarcity}_{g,n,t}$ is the sales-weighted average exposure to sector–commuting zone labor scarcity. The sector–commuting zone-level measure of labor scarcity, $\textit{Zone labor scarcity}_{z,n,t}$, is high when important occupations in sector n and commuting zone z are scarce. In other words, the firm-level measure $\textit{Labor scarcity}_{g,n,t}$ is high when important occupations to production in sector n and commuting zone z, where firm g is present, are scarce.

Growth opportunities. We measure the growth opportunities of a sector using the number of new firms created in this sector in a given year (Adelino, Ma, and Robinson, 2017). We obtain the number of firm creations from the French firm registry (*Répertoire des Entreprises – Fichiers Création*).

our data is that we do not observe the allocation of employees to the firm's different sectors of activity. Hence, to compute the occupational composition of sector n, we cannot precisely distribute employees to the different sectors; instead, we consider all employees of all firms that have as their main sector of activity sector n. This limitation may introduce measurement error correlated with residuals $\eta_{g,n,o,t}$ at the sector of entry level. The precision of this error term may also vary with the sector of origin. Therefore, we double-cluster standard errors to control for correlations within the sector of entry and within the sector of origin.

3.2. Strategy

3.2.1 Empirical model

We test Main Prediction 1, which states that the firm is more likely to "buy" than "build" when it lacks internal resources to operate in the sector of entry, by correlating the mode of entry with the firm's human capital distance to the sector of entry's. We specify the main test as follows:

$$\mathbb{1}(\text{Buy})_{g,n,t} = \lambda_{n,o,t} + \delta \text{Human Capital Distance}_{g,n,t-1} + \beta' X_{g,n,t-1} + \varepsilon_{g,n,t} \tag{6}$$

where the dependent variable $\mathbb{1}(\text{Buy})_{g,n,t}$ is a dummy variable equal to one if the entry of firm g in sector n in year t is made through an acquisition ("buy") and equal to zero if the entry is made organically ("build") (see Section 3.1.2). The main independent variable $\text{Human Capital Distance}_{g,n,t-1}$ is the distance between firm g's human capital and the sector of entry n's human capital in year $t-1$ before the entry (see Section 3.1.3). The vector $X_{g,n,t-1}$ includes other firm characteristics that may influence the decision to build or buy, such as the firm's size, measured by the number of employees, labor productivity, the wage ratio, tangibility of assets, cash holdings and a measure of the diversity of occupations.[24] Definitions of the control variables appear in the appendix table describing the variables.

$\lambda_{n,o,t}$ stands for sector of origin $\times$ entry $\times$ time fixed effects and allows us to compare firms that build and firms that buy within the same sector of entry in the same year and that operate in same sector of origin. The fixed effects strategy captures the endogenous choice of diversification driven by unobservable time-varying synergies between sectors. In addition, this specification also controls for unobservable factors related to the sector of entry, such as fixed costs and barriers to entry (McCardle and Viswanathan, 1994), as well as unobservable factors related to the sector of origin, such as the ability to collateralize assets to access external finance (Rajan and Zingales, 1998). We double-cluster standard errors at the sector of origin and at the sector of entry to control for correlations between firms within the sector of entry and within the sector of origin.

[24] All control variables are scaled by number of employees, as this has the advantage of being additive when aggregating subsidiaries at the group level.

3.2.2 Instrumental variable strategy

Firms may endogenously adjust their workforce composition the year before entry to anticipate a mode of entry. For instance, it may fire redundant employees the year before a buy entry or it may hire employees in key occupations if the firm anticipates a build entry. To address this potential concern, we test the robustness of our main specification to the inclusion of an IV. Following Boehm, Dhingra, and Morrow (2019), we calculate the independent variable HC Distance$_{g,n,t_0}$, which is the HC distance calculated the first year when firm g appears in the sample:

$$HC\ distance_{g,n,t_0} = \sum_i s_{i,g,t_0} \cdot s_{i,n,t_0},$$

where $s_{i,g,t}$ is the share of employees employed in occupation i in firm g and $s_{i,n,t}$ is the share of employees employed in occupation i in sector n.

In addition, we address the concern that firms operating in the same sector may have different production functions and as a result different employment compositions, such that our fixed effects strategy in equation 6 would fail to capture unobservable synergies between the firm's sector of origin and the sector of entry. In this case, synergies would be firm-sector specific. We use the human capital distribution of firms operating in the same sector to construct a sector-level measure of human capital distance. In the spirit of Lee, Mauer, and Xu (2018), we define the firm's weighted HC distance between sectors of origin and entry as follows:

$$HC\ distance^{Sector}_{g,o,n,t} = \sum_o w_{g,o,t} \cdot \left(1 - \sum_i s_{i,o,t} \cdot s_{i,n,t}\right) = \sum_o w_{g,o,t} \cdot HC\ distance_{o,n,t}$$

where $w_{g,o,t}$ the share of sales realized in sector of origin o by firm g in year t, $s_{i,o,t}$ and $s_{i,n,t}$ are the shares of employees employed in occupation i in sectors o and n, respectively. $HCdistance^{Sector}_{g,n,t}$ reflects the fact that firm g's human capital should be close to that of sector n if the sector o in which it operates is close to sector n and if sector o represents a large part of firm g's current sales.

Empirically, we regress these two variables, HC Distance$_{g,n,t_0}$ and HC distance$^{Sector}_{g,o,n,t}$, directly in the reduced-form estimation by relying on OLS estimates. We also specify a 2SLS regression using $HC\ distance^{Sector}_{g,o,n,t}$ as the instrument. We argue that the relevance and the exclusion restrictions are satisfied for these two pseudo IVs. In particular, we observe that HC Distance$_{g,n,t_0}$ and HC distance$^{Sector}_{g,o,n,t}$ are highly correlated with HC Distance$_{g,n,t-1}$. We also

argue that the choice between build or buy is only affected by these instruments through the human capital distance channel taken at $t-1$.

3.2.3 Omitted variables and selection issues

Omitted variables. Our main empirical model specified in equation 6 may suffer from omitted variable bias that drives the choice to build or buy. For instance, the CEO may be the only decision-maker in the choice to build or buy, of the human capital composition, and of firm's scope of activities (Malmendier and Tate, 2005). If it were the case, HC distance would play a very limited role. Ideally, we would like to compare firms that are similar in every aspect except their human capital composition.

Our argument regarding this omitted variable issue is twofold. First, in the baseline model, we limit the role of unobservable variables by constructing a matched sample of building and buying firms, based on the sector of origin, the sector of entry and the firm's size before entry (measured by the quartile of sales in the sector of origin at $t-1$).[25] In the language of the model, this choice is made such firms have comparable c_0 (the firm's core competence in the sector of origin) so that all firms in the sample are large enough to enter a new sector by buying (see Section 2.2).

Second, we check that the correlation between the choice to build or buy and HC distance holds within firms. The introduction of firm fixed effects neutralizes the role of unobservable firm-level factors. However, this specification induces a restriction on our sample, as only firms that make at least two diversifying entries are included in the sample. Moreover, the coefficient is identified for firms that perform at least one buy entry and one build entry during the sample period.

Selection of the sector of entry. A potential limitation of our analysis is that it focuses on firm's choice between build and buy without modeling the firm's decision to diversify in the first place. If the entry and type of entry into the new sector are jointly determined and driven by omitted factors, then our OLS estimates would be biased. From our theoretical analysis, we learn that firms should diversify first in sectors that are closer to their core competence c_0. Indeed, the marginal cost γ to produce in such close sectors is lower, as the firm already has

[25] Specifically, for every buy entry, we only retain build entries within the same (origin, entry, size quartile) triplet.

adequate resources (Section 2.2). Thus, firms should diversify first in the sectors with the lowest human capital distance. We test whether firms indeed diversify in sectors that require human capital close to the firm's core competence. We specify the following test:

$$\mathbb{1}(\text{Diversify})_{g,o,n,t} = \lambda_{g,t} + \lambda_{o,n,t} + \delta \textit{HC Distance}_{g,n,t\text{-}1} + \beta' X_{g,n,o,t-1} + \eta_{g,n,o,t}$$

where $\mathbb{1}(\text{Diversify})_{g,o,n,t}$ is a dummy variable that takes value one if firm g operating in sector o enters sector n in year t, zero otherwise. $\textit{HC Distance}_{g,n,t\text{-}1}$ and the vector of control variables $X_{g,n,o,t-1}$ are those defined in the baseline model 6. $\lambda_{o,n,t}$ is the sector of origin×destination×year fixed effect, and $\lambda_{g,t}$ is a firm×year fixed effect. Because we compute the HC distance between each diversifying firm and every potential sector of entry n and observe the sectors in which a firm diversifies, we can introduce firm×year fixed effects $\lambda_{g,t}$ that capture the firm's likelihood to diversify in a given year. The introduction of $\lambda_{g,t}$ allows us to identify the *direction* of changes in the industry mix – that is, which sectors a firm enters – as opposed to changes in the *number* of industries in which the firm operates.

Selection in diversification. Although we are able to address potential issues related to the selection of the sector of entry, diversifying firms and focused firms may still differ from the outset. From the theoretical analysis, we learn that firms with a high core competence in their sector of origin (large c_0) and specialized resources (small γ) would not diversify. This finding is consistent with the industrial organization literature that finds that market power is an important determinant of a firm's profit within a sector. Note that the decision to build or buy could also be studied in a within-sector setting (i.e., horizontal mergers versus expansion within sector), but we expect that the decision in this case would have very little to do with the degree of specialization of firms' resources – the firm's core competence c_0 likely has the most importance in that context.

In the context of diversification, the parameter γ drives the decision to build or buy. We show that firms with diversified internal human capital build and firms with more specialized human capital but are large enough buy in distant sectors. In the empirical analysis, we do not address the choice of diversification in the first place. However, our analysis sheds light on questions related to *when* firms diversify, *in which* sectors firms diversify, and new to the literature, *how* firms diversify, rather than whether firms diversify. To this end, we restrict the

sample to the universe of firms that diversify at least once in the sample period.

4. Descriptive Statistics

Build or buy? Panel A of Table 1 presents the evolution of build and buy entries between 2005 and 2014. Overall, 1.85% of entries are made by acquisition. This figure is fairly stable over time, although we observe an increase in the number of M&A deals in the second part of the sample from 2009. Buy entries are, on average, larger than build entries. Although sales are higher for build entries than for buy entries in the aggregate, when weighting by entry sales, buy entries increase to an average of X% of total entries between 2004 and 2014.

[Insert Table 1 here]

Human capital distance. Panel A of Table 2 reports summary statistics on our main independent variable, HC Distance. The average diversifying firm has a HC distance of .76 from the sector of entry. HC Distance measures the extent to which a firm employs workers from key occupations for the sector prior to entry. Compared to the maximal distance of 1, this figure suggests that the average firm lacks employees in occupations that are key to operate in the sector of entry. Our baseline measure defined in Section 3.1.3 weights occupations by the number of employees in a given occupation. When we weight HC distance by the occupation wages and the number of hours worked, HC distance remains stable, suggesting that the distribution of wages and hours worked are similar to the distributions of employees by occupation.

In addition, we observe that HC Distance$_{g,n,t_0}$, calculated at t_0, the first year the firm appears in the sample, is slightly larger than the baseline measure, HC Distance$_{\text{g,n,t-1}}$, suggesting that firms' human capital composition evolves over time and becomes slightly closer to the sector of entry's human capital composition. By contrast, we find that the HC sector distance, HC Distance$^{\text{Sector}}_{\text{g,o,n,t-1}}$, is smaller than the firm-level HC distance, suggesting that the human capital composition has a firm-specific component.

Other variables. Distributions of the control variables we include in our baseline specification (equation (6)) are also reported in Panel A of Table 2. Firms employ on average X employees, produce approximately €60,000 in value added per worker, own €40,000 in tangible assets

per employee, and hold €20,000 in cash per worker. Of diversifying firms, 45% were already diversified before the entry into the new sector. Only 2% of them include a listed subsidiary in the business group.

Firms report on average €4.8 million in sales in the sector of entry in the first year, with a very large dispersion around the mean. Sales in the new sector represent on average 116% of sales in the sector of origin the year before diversification; 24% of the diversification occurs in an upstream sector.[26] Only 36% of the diversification in a different 3-digit sector occurs in the same 1-digit industry, and 24% occurs in the same 2-digit industry. In addition, only 4% of the diversification occurs in a different region, and 7% occurs in a different department. These stylized facts suggests that the average firm diversifies in a distant sector from its sector of origin but in the same geographic area.

[Insert Table 2 here]

5. Main Results

5.1. Human capital and corporate diversification

We test whether firms are more likely to buy when their human capital is distant from the sector of entry (Main Prediction 1). The dependent variable $\mathbb{1}(Buy)_{g,n,t}$ is a dummy variable equal to one if firm g enters a new sector n in year t through the acquisition of an incumbent firm and zero if it enters by building on its own resources. The main independent variable HC $\text{Distance}_{\text{g,n,t-1}}$ is high when the diversifying firm employs few employees in key occupations for the sector.

Table 4 reports the results with different sets of fixed effects. The specifications in columns 1 and 2 include interacted sector of origin × sector of entry × year fixed effects. These specifications allow us to compare firms that operate in the same sector of origin and enter the same new sector in the same year, while controlling for potential unobservable synergies between the sector of origin and that of entry. The specification in column 2 includes firm-level control variables: size, total cash holdings, tangible assets, and value added, with the latter variables being scaled by the number of workers. The specification in column 2 is our baseline specification. In

[26] A sector of entry n is said to be upstream of a sector of origin o if more than 5% of the input used by firms in sector o comes from sector n.

all models, standard errors are double-clustered at the sector of origin and sector of entry.

We show that the firm's human capital distance to the sector of entry is positively correlated with the probability of entering by acquisition. A one-standard-deviation increase (22 percentage points) in human capital distance is associated with a 0.4 percentage-point increase in the likelihood of entering by buying (columns 1 and 2). This relationship is sizable, equal to 20% of the unconditional probability of buying, and significant at the 5% level. The point estimate is relatively unchanged when we add control variables in column 2, suggesting that the control variables are mostly uncorrelated with the human capital distance.

In columns 3 to 4, we test the robustness of our main finding to different combinations of fixed effects. In column 3, we include interacted sector of origin $\times$ entry $\times$ size quartile fixed effects to compare firms of the same size in their sector of origin. In column 4, we compare firms that initially operate in the same *main* sector of origin and the same *second* sector of origin and enter the same sector. This specification accounts for possible additional unobservable synergies with the sector of entry and for the fact that some business groups are already diversified. The positive relationship between the likelihood of buying and human capital distance is robust to these alternative combinations of fixed effects.

In column 5, we include firm fixed effects in addition to year dummies to test the robustness of our results to unobservable firm-level determinants of the choice to build or buy. We also include interacted sector of origin $\times$ entry fixed effects to maintain the comparison of firms in the same pair of sectors. This specification limits the role of omitted variables and addresses concerns related to the selection of firms into a given mode of entry (see Section 3.2.3). Within firms, the relationship between buy entries and human capital distance remains significant at the 5% level. However, within firms, the point estimate is half of that in specifications relying on the cross-sectional variation between firms (columns 1 to 4), suggesting that unobservable firm-level factors co-determine the choice to build or buy in addition to the firm's human capital. The evidence shows that firms' human capital is undoubtedly an important determinant of the mode of entry into sectors. We determine that firm human capital accounts for approximately 50% of the variation within firms.[27]

Importantly, in columns 2 to 4, we control for the firm's human capital diversity of occupa-

[27]Note that in the within-firms specification (column 5), the coefficients of firm size, value added per employee and the average number of occupations have a different sign than in the other columns. Our interpretation is that firms that diversify multiple times tend to buy at a stage of their life-cycle in which they are relatively smaller and less profitable than when they build.

tions and firm's size measured by the number of employees. Our main finding is not affected by the introduction of these control variables in addition to the other variables. In particular, we find that larger firms are more likely to buy than to build. This is consistent with the comparative statics of the model derived in Section 2. In line with prior literature on the determinants of M&A deals, we find that financially unconstrained firms (cash-rich firms) and high-paying firms are significantly more likely to buy than to build. Interestingly, the results also show that firms that employ a relatively diverse set of occupations are more likely to perform acquisitions (columns 2 to 4). This relationship is negative within firms (column 5), suggesting that firms that become marginally more diverse and smaller are in contrast less likely to buy.

In conclusion, the empirical analysis presented in 4 supports the model's main predictions. Firms buy rather than build when their human capital is distant from, and therefore less adapted to operate in, the sector of entry. We show that this result holds in the cross section of firms, as well as within firms over time. It is also robust to various combinations of sector fixed effects and to the introduction of other determinants of M&As.

[Insert Table 4 here]

In unreported results, we interact the human capital distance with different measures of firm size and access to capital (number of employees, an indicator of whether firms are already diversified, cash holdings per employee and an indicator of whether the firm is publicly listed). Consistent with the model, we find that the role of human capital distance is less pronounced for larger firms and for firms that are already diversified. This finding suggests that larger firms (low c_0) are better able to build in more distant sectors.

5.2. Robustness

Table A4 tests the robustness of our main finding to alternative measures of human capital (Panel A), the endogeneity of firms' human capital using an IV strategy (Panel B), and the size of the entry into the new sector (panel C).

Alternative measures of human capital. Panel A, column 1 reports the results of our baseline specification (equation (6). It correlates the mode of entry with the firm's human capital distance to the sector of entry. HC distance is weighted by the number of employees in

each occupation, with a larger weight being assigned to quantitatively important occupations. In contrast, occupations with few individuals have a lower weight (e.g., CEOs and other high-level executives). In columns 2 and 3, we weight occupations by wages and the collective number of hours worked.

An alternative explanation for our main finding is that firms may anticipate the mode of entry by adjusting the composition of their human capital several years before entry. In this case, firms' human capital would be endogenous to the mode of entry, and our interpretation would be biased. To overcome this issue, we regress the choice to buy rather than build on the maximum lagged value of human capital distance (column 4). We show that a firm's maximum lagged human capital distance is still positively and significantly related to the likelihood of diversifying by acquisition. However, the point estimates on the maximum lagged human capital distance are significantly lower that in our baseline estimation, suggesting that firms adjust their workforce composition during the years preceding entry, although without perfectly anticipating the mode of entry.

Industry average of human capital. In panel B, we test the robustness of our main result to further endogeneity concerns. First, we regress the choice to buy on the sector of origin's human capital distance to the sector of entry. The instrument exploits the average human capital similarity of firms in the sector. In columns 1 and 2, we regress the firm's choice to buy directly on the sector HC distance. We replace the combination of origin $\times$ entry $\times$ year fixed effects with origin + entry + year fixed effects to have enough variation to estimate the coefficient of sector HC. We still find that the sector human capital distance is positively related to the decision to buy rather than build.

Next, we perform a 2SLS estimation. In the first stage, we regress the sector-level human capital distance on the firm-level human capital distance (column 3). The point estimate is .54 and highly significant. The second-stage results are reported in column 4. The relationship between human capital distance and the likelihood of diversifying by acquisition is still positive and significant. The point estimate is larger than in the reduced-form instrumentation strategy reported in column 2, suggesting that a significant part of the variation in the firm's human capital is sector-specific.

Size of the entry. In panel C, we test the robustness of our preferred specification to the new sector's size of entry. The entry size is measured as the ratio of sales realized in the new sector relative to sales realized in the other sectors where the firm operates.

In column 1, we do not impose any threshold of entry sales. We still find a positive relationship between acquisitions and the HC distance to the sector of entry. However, the relationship is significant only at the 10% level, and the magnitude of the effect is significantly smaller. In column 2, we impose a threshold of entry sales equal to 1%. Column 3 is our baseline model, where entry sales represent at least 5% of firm's portfolio of activities. In column 4, we restrict the sample to entry sales that represent at least 10% of the firm's total sales. We find that the relationship between buy entries and human capital distance becomes economically and statistically stronger as we impose a higher threshold for entry sales. This finding suggests that the choice to buy or build is more sensitive to the firm's human capital when the entry into the new sector is meaningful at the scale of the diversifying firm. In other words, when the entry is small, possibly driven by experimental motives, buying an incumbent firm is less of an option, and workforce composition is not as crucial as in larger entries. We confirm this intuition in column 5 by performing a weighted least squares (WLS) regression, where observations are weighted by the share of entry sales.

[Insert Table A4 here]

5.3. The role of product market and geographic distance

Firms may diversify by acquisition in more distant sectors from the sector of origin and may instead enter close sectors by building, irrespective of labor considerations. Firms may also behave differently when they enter a vertically integrated sector or when they enter an unrelated sector. In addition, the geographic distance between the initial firm's location and the effective location of the new activities may play an important role, with more geographicly distant firms lacking the local resources to enter a sector in a new geographic area by building. In Table A3, we investigate these effects.

First, we interact the human capital distance with a simple measure of sectoral distance based on the French nomenclature of sectors. The sector of origin and the sector of entry are close to one another if they are classified under the same 1-(2-)digit industry code, respectively.[28]

[28]Hoberg and Phillips (2010) develop a measure based on textual analysis of 10k filings that allows one to

Furthermore, the approach in columns 1 to 4 consists of replacing the sector of origin × entry × year fixed effects with the combination sector of origin × year + sector of entry × year fixed effects to be able to estimate the effect of sectoral distance on the decision to build or buy.[29]

The results show that firms tend to buy in the same industry rather than in more distant industries (columns 1 and 2). However, the effect is eliminated when the firm's human capital is distant from that needed in the sector of entry that is itself in the same industry as the sector of origin. In other words, firms are more likely to buy than build in two situations: when their human capital is distant from that of a distant sector from the sector of origin and when their human capital is adapted to a sector that is close to the sector of origin.[30] This finding suggests that diversification in distant sectors amplifies the importance of human capital in the decision to build or buy.

Second, in columns 3 and 4, we test the effect of vertical integration on the decision to build or buy. Firms may be willing to enter upstream sectors through the acquisition of a supplier or by transferring human capital along the production chain. We identify a vertical link between a firm and the sector of entry when more than 5% and 10% of the inputs used by the sector of origin come from the sector of entry.[31] We find that firms are more likely to buy in an upstream sector. However, the effect is eliminates when human capital is distant from the upstream sector of entry.

Third, we investigate the role of the physical distance between the firm and the new activities' location in columns 5 and 6. geographic diversification could induce firms to buy rather than build if tapping a new market requires a physical presence, e.g., to build local customer capital or if a physical presence is a proxy for better knowledge of the local market. Moreover, if potential targets are physically close, then it is less expensive to build to enter a geographicly distant market. To test the effect of geographic distance, we flag entry into new geographic

precisely assess whether firms are close competitors in the product market space. The authors show that this measure sharply improves upon widely used industry classifications in explaining many different firm-specific decisions such as firm profitability, dividends and Tobin's Q. Unfortunately, the measure is only available for US publicly listed firms.

[29]Note that we cannot estimate these effects in our baseline specification, which includes sector of origin × entry × year fixed effects to control for the endogenous matching between the main sector of origin and the main sector of entry.

[30]Note that this second effect is consistent with Rhodes-Kropf and Robinson (2008) and Hoberg and Phillips (2010), who find that firms that make similar products tend to merge.

[31]We construct the measure of vertical integration following Fan and Goyal (2006). We use a 2013 input-output matrix for France compiled by the French Bureau of Statistics (Insee). The dimension of the matrix is 138 industries by 138 industries. Note that the regression excludes observations for which the industry of origin and the industry of entry are the same, the vertical link variable not being defined in this case.

areas at the *Department* and *Region* levels.[32] The evidence reveals that 58% and 23% of buy and build entries, respectively, occur in a new department. In line with this, we find that entries in new area are associated with buy entries. However, the interaction of geographic distance with human capital is not significantly different from zero. The evidence suggests that the role of geographic distance in a firm's decision to build or buy is distinct from that of human capital.

[Insert Table A3 here]

5.4. Human capital and post-entry performance

In the previous section, we show that firms that do not employ the right set of employees tend to diversify by acquisition. In this section, we provide evidence that not having the right human capital to operate, thus having a high HC distance to the sector of entry's, is associated with worse ex post performance. Table **??** reports the results. In this table, we focus on building firms and measure performance at the entity level.

The evidence shows that a one-standard-deviation increase in the HC distance is associated with a reduction in entry sales of 15% (column 1). Not having the right human capital also has long-term consequences. Firms with a distant human capital to the sector of entry's sell 20% less over three years than other buying firms coming from the same sector of origin entering the same sector in the same year (column 3). Moreover, having distant human capital is associated with a lower probability of surviving in the sector (column 2). These firms are 35% less likely to survive three years after entry. Surprisingly, the finding in column 4 suggests that firms with a high HC distance do not significantly employ more people over the three years following entry. The net effect on employment is not significantly different from zero. However, we find that these firms with an unadapted human capital significantly reduce their HC distance to the sector of entry within the three next years. A one-standard-deviation increase in HC distance in the year before entry is associated with a 9.6 p.p. decrease in HC distance after three years.

Overall, we show that firms that build despite their human capital not being adapted to the sector of entry sell less and are less likely to survive after three years. However, the evidence shows that these firms adjust their human capital following entry and hire employees in the key

[32] France has 25 régions and 96 départements in the sample period. Note that we are able to identify whether a firm begins to operate in a new geographic area in two cases: first, where a new subsidiary located in the new area is acquired and, second, when the entering firm opens a new establishment in the new area in year t. In contrast, we are not able to identify entries in new geographic areas that are not associated with the creation or the acquisition of an establishment.

occupations for the sector of entry, thus reducing their HC distance to the sector of entry over time. These findings validate our measure of human capital.

[Insert Table **??** here]

6. Channels

Thus far, we have shown that firms buy when their human capital is distant from the sector of entry's. The next question is, given the costs associated with an acquisition, why do firms not hire new employees instead of buying an existing firm? The model predicts that firms are more likely to buy when the cost of hiring specialized workers are high and when growth opportunities in the sector of entry are important. We then test these predictions in the cross-section of building and buying firms. We interact our main independent variable HC Distance$_{g,n,t\text{-}1}$ with two sources of plausible exogenous variation. First, we test that the effects of a firm's human capital when some key occupations are scarce resources in the economy. Second, we test the effects of a firm's human capital in the presence of growth opportunities.

6.1. Cross-sectional variation

Labor scarcity$_{g,n,t\text{-}1}$is the sales-weighted average exposure to the sector–commuting zone labor scarcity (see Section 3.1.4 for more details). It reflects the extent to which occupations in the commuting zone where the firm operates are scarce. The empirical model with HC distance interacted with Labor scarcity becomes:

$$\begin{aligned}\mathbb{1}(\text{Buy})_{g,n,t} = \lambda_o + \lambda_n + \lambda_t + \delta_1 \text{HC Distance}_{g,n,t\text{-}1} + \delta_2 \text{Labor Scarcity}_{g,n,t\text{-}1} \\ + \delta_3 \text{HC Distance}_{g,n,t\text{-}1} \times \text{Labor Scarcity}_{g,n,t\text{-}1} + \beta' X_{g,n,o,t-1} + \varepsilon_{g,n,o,t}\end{aligned} \tag{7}$$

We expect $\delta_3 > 0$, suggesting that it is even more valuable to diversify by buying when the HC Distance$_{g,n,t\text{-}1}$ to the sector of entry is high and when the key occupations to operate in the sector of entry are scarce.

We perform a similar cross-sectional test with growth opportunities defined as the number

of new firm creations in the sector of entry, which is given by the following equation:

$$\begin{aligned}\mathbb{1}(\text{Buy})_{g,n,t} = \lambda_o + \lambda_n + \lambda_t + \delta_1 \text{HC distance}_{g,n,t\text{-}1} + \delta_2 \textit{New firms}_{n,t-1} \\ + \delta_3 \text{HC distance}_{g,n,t_0} \times \textit{New firms}_{n,t-1} + \beta' X_{g,n,o,t-1} + \varepsilon_{g,n,o,t}\end{aligned} \tag{8}$$

We expect $\delta_3 < 0$, as in presence of growth opportunities, firms should be relatively less likely to buy in sectors distant from their core competence. From the theoretical analysis, we learn that firms buy in sectors relatively closer to their core competence when they observe important growth opportunities in these sectors. Firms give up profits in distant sectors by making opportunistic acquisitions in closer sectors where doing so become relatively more productive and hence the investment relatively more profitable (a is large in the model).[33] As a result, in presence of growth opportunities in close sectors, it becomes relatively more attractive to buy in sectors distant from the firm's core competence.

Note that equations 7 and 8 include a combination of sector of origin + sector of entry + year of entry fixed effects. There are two reasons for this deviation from our baseline specification (equation 6). First, the variable $\textit{New firms}_{n,t-1}$ varies at the sector of entry-year level, which does not allows us to include sector of entry×year fixed effects. Second, although the variable $\textit{Labor scarcity}_{g,n,t-1}$ varies at the sector of origin-entry-year level, it imposes substantial constraints on our analysis. First, we have data on scarce occupations from 2010 to 2014 only, such that we can only test the effect on a limited sample of three years of data. In addition, the list of scarce occupations does not change substantially over time, which means that we also cannot include interacted sector×year fixed effects in this specification.

6.2. The effects of labor scarcity and growth opportunities

Table **??** reports the results. In columns 1 and 2, we interact HC distance with labor scarcity following equation 7. In columns 3 and 4, we interact HC distance with the number of new firms created in the sector of entry in a given year, following equation 8. For clarity of interpretation, we split the continuous variable labor scarcity into dummy variables that represent terciles of the respective distribution.

Columns 1 and 2 of Table **??** show that the point estimates on the second and third terciles

[33] At the extreme (very large a), firms should always diversify by buying, irrespective of the HC distance and their internal resources.

of labor scarcity interacted with HC distance are positively associated with the decision to buy. The evidence shows that a firm is 1.3 to 2.8 percent more likely to buy when it does not already employ the key occupations to operate in the sector of entry and when these occupations are scarce in the commuting zone where the firm is present. These findings hold both with and without control variables.

Columns 3 and 4 of Table **??** show that firms are significantly more likely to buy in sectors where new firms are created. However, this effect disappears when firms' human capital is distant from the that of the sector of entry. The evidence shows that firms buy relatively more in sectors close to their core competence when they experience growth opportunities and derive more profit than they did formerly. Indeed, buying firms corner a higher share of the target companies in presence of growth opportunities, such that it becomes relatively more attractive to buy in these closer sectors. At the extreme, this finding suggests that firms opportunistically buy in dynamic sectors irrespective of their human capital.

Overall, the findings in this section are consistent with the prediction that human capital determines firms' decision to build or buy and is especially acute in the presence of scarce occupations but less important in presence of growth opportunities (see Section 2).

[Insert Table **??** here]

7. Selection of the Sector of Entry

A potential limitation of our analysis is that we focus on the tradeoff of build versus buy without eliciting the decision to diversify in the first place. If the entry and type of entry into the new sector are jointly determined and driven by omitted factors, then the OLS estimates may be biased. Our aim in this section is to understand why firms diversify into certain sectors and whether human capital distance explains this choice.[34]

Our approach consists of computing the firm's human capital distance to every sector and then correlating this variable to the choice of sector the firm actually makes. Thus, the

[34] A classic solution to address selection issues in diversifying choices is to find an IV that affects the choice of sector but is plausibly orthogonal to the main dependent variable. For instance, Tate and Yang (2016a) uses Tobin's Q as an instrument in a two-stage Heckman selection model. Applied to our context, we would need an instrument that affects the choice to diversify but not the type of entry. In addition, this instrument would need to be sector-specific because we would have to instrument not only for the choice to build or buy but also for the decision to diversify in a given sector. Finding such an instrument appears to be a difficult task; therefore, we choose to focus on specific scenarios in which such selection issues are likely to arise.

dependent variable in Table 3 is a dummy variable that indicates whether firm g enters sector n in year t, as opposed to another sector.[35]

Our results show that firms are less likely to diversify in sectors in which their human capital is distant from what is needed to operate. A one-standard-deviation increase in a firm's human capital distance to the sector's is associated with a 0.23 p.p. decrease in the probability of diversifying in this sector. In other words, firms are more likely to diversify in sectors to which their human capital is adapted. The relationship holds when we compare firms in the same sectors of origin and entry in the same year (column 1) and when we compare the same firm over time (column 2). Finally, by examining a given firm and a given year, we again find that firms are significantly less likely to diversify in sectors that are more distant from their core competence.

[Insert Table 3 here]

8. Conclusion

How do firms become conglomerates? Why do some firms enter a new sector by acquiring an existing company ("buy"), while others do so using their existing resources ("build")? This paper jointly studies the role of productive capabilities in determining the firm's scope of activity and the mode of entry into new sectors. When a firm buys to enter a new sector, it has to incur both the costs of acquiring and restructuring the target, but it also secures access to the target's productive resources. When a firm builds on its existing resources to enter a new sector, it must pay the adjustment costs needed to acquire an adapted set of capabilities.

We show that the vast majority of firms enter new sectors by building on their preexisting resources. Only 2% of firms (10% weighted by entry sales) diversify by buying, but these firms represent We focus on the role of human capital to explain patterns of corporate diversification and construct a firm-level measure of human capital based on the occupational composition of firms' workforces. We show theoretically and empirically that firms are more likely to build or grow organically. However, firms are more likely to buy in sectors that require skills that they lack, i.e., when their human capital is distant from the sector of entry's. This result suggests that labor adjustment costs determine choice of the sector of entry and the choice to build or

[35]Note that this specification multiplies the number of observations by x, the number of sectors in the economy.

buy.

We use an IV strategy to ensure that our results are not driven by the endogenous composition of human capital that may arise from the anticipation of the mode of entry. Our results also are robust to alternative definitions of our human capital distance and different thresholds of reported sales in the sector of entry.

We investigate two channels that drive the positive relationship between human capital distance and entry by buying. First, firms are even more likely to buy when key employees for the sector of entry are scarce resources on the external job market. Second, firms are relatively more likely to buy in sectors characterized by growth opportunities regardless of the human capital distance. These two findings in the cross-section of labor and product markets suggest that firm's internal resources are driving firms' decision to build or buy. Thereby, we contribute to the literature on corporate diversification by showing that both internal resources and the cost of accessing external resources play a role in explaining how firms diversify.

References

Abowd, John M and Francis Kramarz. 2003. "The costs of hiring and separations." *Labour Economics* 10 (5):499–530.

Adelino, Manuel, Song Ma, and David Robinson. 2017. "Firm age, investment opportunities, and job creation." *The Journal of Finance* 72 (3):999–1038.

Agrawal, Ashwini K and David A Matsa. 2013. "Labor unemployment risk and corporate financing decisions." *Journal of Financial Economics* 108 (2):449–470.

Baghai, Ramin, Rui Silva, Viktor Thell, and Vikrant Vig. 2018. "Talent in distressed firms: Investigating the labor costs of financial distress." *Working Paper* .

Bai, John Jianqiu, Douglas J Fairhurst, and Matthew Serfling. 2018. "Employment protection, investment, and firm growth." Working paper.

Berger, Philip G and Eli Ofek. 1995. "Diversification's effect on firm value." *Journal of financial economics* 37 (1):39–65.

Bernard, Andrew B, Emily J Blanchard, Ilke Van Beveren, and Hylke Vandenbussche. 2018. "Carry-along trade." *The Review of Economic Studies* 86 (2):526–563.

Bernard, Andrew B, J Bradford Jensen, Stephen J Redding, and Peter K Schott. 2007. "Firms in international trade." *Journal of Economic perspectives* 21 (3):105–130.

Bernard, Andrew B, Stephen J Redding, and Peter K Schott. 2010. "Multiple-product firms and product switching." *American Economic Review* 100 (1):70–97.

Betton, Sandra, B Espen Eckbo, and Karin S Thorburn. 2008. "Corporate takeovers." In *Handbook of empirical corporate finance*. Elsevier, 291–429.

Blatter, Marc, Samuel Muehlemann, and Samuel Schenker. 2012. "The costs of hiring skilled workers." *European Economic Review* 56 (1):20–35.

Boehm, Johannes, Swati Dhingra, and John Morrow. 2019. "The comparative advantage of firms." *Working Paper* .

Broda, Christian and David E Weinstein. 2010. "Product creation and destruction: Evidence and price implications." *American Economic Review* 100 (3):691–723.

Campa, Jose Manuel and Simi Kedia. 2002. "Explaining the diversification discount." *Journal of Finance* 57 (4):1731–1762.

Cestone, Giacinta, Chiara Fumagalli, Francis Kramarz, and Giovanni Pica. 2016. "Insurance between firms: the role of internal labor markets." Tech. rep., Working Paper.

Chandler, Alfred D. 1992. "Organizational capabilities and the economic history of the industrial enterprise." *Journal of Economic Perspectives* 6 (3):79–100.

Chen, Deqiu, Huasheng Gao, and Yujing Ma. 2018. "Human capital driven acquisition: Evidence from the inevitable disclosure doctrine." *Working Paper* .

Clara, Nuno, Alexandre Corhay, and Howard Kung. 2020. "Strategic product piversity." *Working Paper* .

D'Acunto, Francesco, Geoffrey Tate, and Liu Yang. 2018. "Entrepreneurial teams: Diversity of experience and firm growth." *Unpublished Working Paper* .

Dessaint, Olivier, Andrey Golubov, and Paolo Volpin. 2017. "Employment protection and takeovers." *Journal of Financial Economics* 125 (2):369–388.

Donangelo, Andres. 2014. "Labor mobility: Implications for asset pricing." *Journal of Finance* 69 (3):1321–1346.

Eckel, Carsten and J Peter Neary. 2010. "Multi-product firms and flexible manufacturing in the global economy." *The Review of Economic Studies* 77 (1):188–217.

Elsas, Ralf, Mark J Flannery, and Jon A Garfinkel. 2014. "Financing major investments: Information about capital structure decisions." *Review of Finance* 18 (4):1341–1386.

Fan, Joseph P. H. and Vidhan Goyal. 2006. "On the patterns and wealth effects of vertical mergers." *Journal of Business* 79 (2):877–902.

Giroud, Xavier and Holger M Mueller. 2015. "Capital and labor reallocation within firms." *Journal of Finance* 70 (4):1767–1804.

Hoberg, Gerard and Gordon Phillips. 2010. "Product market synergies and competition in mergers and acquisitions: A text-based analysis." *Review of Financial Studies* 23 (10):3773–3811.

———. 2018. "Conglomerate industry choice and product language." *Management Science* 64 (8):3735–3755.

Kuehn, Lars-Alexander, Mikhail Simutin, and Jessie Jiaxu Wang. 2017. "A labor capital asset pricing model." *Journal of Finance* 72 (5):2131–2178.

Lagaras, Spyridon. 2017. "Corporate takeovers and labor restructuring." *Working Paper* .

Lamont, Owen. 1997. "Cash flow and investment: Evidence from internal capital markets." *The Journal of Finance* 52 (1):83–109.

Lang, Larry HP and Rene M Stulz. 1994. "Tobin's q, corporate diversification, and firm performance." *Journal of political economy* 102 (6):1248–1280.

Lee, Kyeong Hun, David C. Mauer, and Emma Qianying Xu. 2018. "Human capital relatedness and mergers and acquisitions." *Journal of Financial Economics* 129 (1):111–135.

Ma, Wenting, Paige Ouimet, and Elena Simintzi. 2019. "Mergers and acquisitions, technological change and inequality." *Working Paper* .

Maksimovic, Vojislav and Gordon Phillips. 2002. "Do conglomerate firms allocate resources inefficiently across industries? Theory and evidence." *Journal of Finance* 57 (2):721–767.

Maksimovic, Vojislav, Gordon Phillips, and Nagpurnanand R Prabhala. 2011. "Post-merger restructuring and the boundaries of the firm." *Journal of Financial Economics* 102 (2):317–343.

Maksimovic, Vojislav and Gordon M Phillips. 2013. "Conglomerate firms, internal capital markets, and the theory of the firm." *Annual Review of Financial Economics* 5 (1):225–244.

Malmendier, Ulrike and Geoffrey Tate. 2005. "CEO overconfidence and corporate investment." *Journal of Finance* 60 (6):2661–2700.

———. 2008. "Who makes acquisitions? CEO overconfidence and the market's reaction." *Journal of financial Economics* 89 (1):20–43.

Matsa, David A. 2010. "Capital structure as a strategic variable: evidence from collective bargaining." *Journal of Finance* 65 (3):1197–1232.

McCardle, Kevin F and S Viswanathan. 1994. "The direct entry versus takeover decision and stock price performance around takeovers." *Journal of Business* :1–43.

Merz, Monika and Eran Yashiv. 2007. "Labor and the market value of the firm." *American Economic Review* 97 (4):1419–1431.

Ouimet, Paige Parker and Rebecca Zarutskie. 2016. "Acquiring labor." Tech. rep., Working Paper.

Penrose, Edith. 1955. "Limits to the growth and size of firms." *American Economic Review* 45 (2):531–543.

Phillips, Gordon M and Alexei Zhdanov. 2013. "R&D and the Incentives from Merger and Acquisition Activity." *The Review of Financial Studies* 26 (1):34–78.

Rajan, Raghuram, Henri Servaes, and Luigi Zingales. 2000. "The cost of diversity: The diversification discount and inefficient investment." *Journal of Finance* 55 (1):35–80.

Rajan, Raghuram and Luigi Zingales. 1998. "Financial Dependence and Growth." *American Economic Review* 88 (3):559–86.

Rhodes-Kropf, Matthew and David T. Robinson. 2008. "The market for mergers and the boundaries of the firm." *Journal of Finance* 63 (3):1169–1211.

Romanski, Hilton. 2017. "Cisco's 200th Acquisition - A Tradition of Advancement, Disruption and Growth." *Cisco Blog* .

Scharfstein, David S. 1998. "The dark side of internal capital markets II: Evidence from diversified conglomerates." Tech. rep., National Bureau of Economic Research.

Scharfstein, David S and Jeremy C Stein. 2000. "The dark side of internal capital markets: Divisional rent-seeking and inefficient investment." *Journal of Finance* 55 (6):2537–2564.

Schoar, Antoinette. 2002. "Effects of corporate diversification on productivity." *The Journal of Finance* 57 (6):2379–2403.

Serfling, Matthew. 2016. "Firing costs and capital structure decisions." *Journal of Finance* 71 (5):2239–2286.

Shin, Hyun-Han and René M Stulz. 1998. "Are internal capital markets efficient?" *The Quarterly Journal of Economics* 113 (2):531–552.

Simintzi, Elena, Vikrant Vig, and Paolo Volpin. 2014. "Labor protection and leverage." *Review of Financial Studies* 28 (2):561–591.

Stein, Jeremy C. 1997. "Internal capital markets and the competition for corporate resources." *The Journal of Finance* 52 (1):111–133.

Tate, Geoffrey and Liu Yang. 2016a. "The bright side of corporate diversification: evidence from internal labor markets." *Review of Financial Studies* 28 (8):2203–2249.

Tate, Geoffrey A and Liu Yang. 2016b. "The human factor in acquisitions: cross-industry labor mobility and corporate diversification." *Working Paper* .

Wysocki, Bernard Jr. 1997. "Many mergers driven by search for fresh talent." *Wall Street Journal* .

Xu, Sheng-Jun. 2018. "Skilled labor supply and corporate investment: Evidence from the H-1B visa program." *Working Paper* .

Yip, George S. 1982. "Diversification entry: Internal development versus acquisition." *Strategic Management Journal* 3 (4):331–345.

Figures and Tables

Table 1. Evolution of Build and Buy Entries

This table reports the number and relative frequency of buy entries and build entries. A firm "buys" when it enters a new sector through M&A. *Source:* SDC Platinum, BvD Zephyr, EAE/VAC survey, tax files, ownership links dataset, and matched employer-employee dataset. *Sample:* Firms that enter a new sector during the 2005-2014 period.

Panel A. Build and buy entries by year

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Total
Build (number)	12,548	10,576	11,460	15,548	12,952	18,019	11,973	11,544	11,320	11,245	127,185
Buy (number)	165	246	216	236	165	200	282	446	200	259	2,415
Build/buy (%, equally-weighted)	1.30	2.27	1.85	1.50	1.26	1.10	2.30	3.72	1.74	2.25	1.93
Build/buy (%, entry sales-weighted)	11.73	19.81	15.11	9.73	6.05	7.04	19.46	36.96	19.74	20.80	16.64

Panel B. Build and buy entries by industry classification level

Industry level:	5 digits	4 digits	3 digits	2 digits	1 digit
Build (number)	157,792	144,730	127,185	96,750	75,528
Buy (number)	2,650	2,533	2,415	1,811	1,088
Buy (%, equally-weighted)	1.65	1.72	1.86	1.84	1.42
Buy (%, entry sales-weighted)	10.79	12.78	15.68	21.02	20.75

Table 2. Summary Statistics

Panel A reports descriptive statistics for firms that enter a new sector n, either by buying or building. An observation is at the level of a firm-sector-time triplet (g, n, t). Panel B displays descriptive statistics for all (g, n, t) triplets, whether firm g enters sector n or not (only firms present in panel A are included). A firm "buys" when it enters a new sector through M&A. A firm "builds" when it enters a new sector through one of its subsidiaries. Sectors refer to an industry at the 3-digit level of the French classification of industries. Descriptions of the variables are reported in Appendix B. *Source:* SDC Platinum, BvD Zephyr, EAE/VAC survey, tax files, ownership links dataset, and matched employer-employee dataset. *Sample:* Firms that enter a new sector and realize at least 5% of their sales in the new sector in the 2005-2014 period.

Panel A. Distribution of firm characteristics

	N	Mean	Std Dev.	5^{th}	25^{th}	Median	75^{th}	95^{th}
$1(\text{Buy})_{g,n,t}$	32,936	0.02	0.13	0.00	0.00	0.00	0.00	0.00
$\text{HC Distance}_{g,n,t\text{-}1}$	32,936	0.76	0.22	0.26	0.66	0.84	0.93	0.99
$\text{HC Distance}_{g,n,t\text{-}1}$ (wages)	32,936	0.78	0.21	0.32	0.68	0.85	0.93	0.99
$\text{HC Distance}_{g,n,t\text{-}1}$ (hours)	32,935	0.76	0.22	0.26	0.66	0.84	0.93	0.99
$\text{HC Distance}_{g,n,t0}$	30,517	0.80	0.21	0.33	0.72	0.87	0.96	1.00
$\text{HC Distance}^{Sector}_{o,n,t\text{-}1}$	32,934	0.60	0.23	0.14	0.46	0.65	0.78	0.91
$\text{Value added/N. Employees}_{g,t\text{-}1}$	32,936	0.06	0.04	0.02	0.04	0.05	0.07	0.12
$\text{N. Occupations/N. Employees}_{g,t\text{-}1}$	32,936	0.40	0.28	0.08	0.20	0.33	0.51	1.00
$\log(\text{N. Employees})_{g,t\text{-}1}$	32,936	3.63	1.21	1.61	3.00	3.61	4.29	5.67
$\text{Tangible Assets/N. Employees}_{g,t\text{-}1}$	32,936	0.04	0.07	0.00	0.01	0.02	0.05	0.15
$\text{Total wages/N. Employees}_{g,t\text{-}1}$	32,936	0.03	0.01	0.02	0.02	0.03	0.03	0.05
$\text{Cash/N. Employees}_{g,t\text{-}1}$	32,936	0.02	0.04	0.00	0.00	0.01	0.03	0.09
$\text{Entry sales (M€)}_{g,n,t}$	32,936	4.80	26.64	0.10	0.38	1.02	2.94	14.97
$\text{Entry sales/Lagged total sales}_{g,n,t}$	32,936	0.33	0.55	0.06	0.09	0.16	0.35	0.96
$1(\text{Diversified})_{g,t\text{-}1}$	32,936	0.45	0.50	0.00	0.00	0.00	1.00	1.00
$1(\text{Public})_{g,t\text{-}1}$	32,936	0.02	0.14	0.00	0.00	0.00	0.00	0.00
$1(\text{Vertical link } >5\%)_{o,n}$	32,936	0.24	0.42	0.00	0.00	0.00	0.00	1.00
$1(\text{Same 1-digit Industry})_{o,n}$	32,936	0.36	0.48	0.00	0.00	0.00	1.00	1.00
$1(\text{Same 2-digit Industry})_{o,n}$	32,936	0.19	0.39	0.00	0.00	0.00	0.00	1.00
$1(\text{New department})_{g,t}$	32,828	0.07	0.25	0.00	0.00	0.00	0.00	1.00
$1(\text{New region})_{g,t}$	32,828	0.04	0.18	0.00	0.00	0.00	0.00	0.00
$\text{Labor scarcity}_{g,n,t\text{-}1}$	9,049	0.18	0.16	0.01	0.06	0.13	0.25	0.50
$\text{N. New firms}_{n,t\text{-}1}$	32,936	2960.15	4026.19	81.00	585.00	1298.00	3074.00	13076.00

Panel B. Selection of sector of entry

	N	Mean	Std Dev.	5^{th}	25^{th}	Median	75^{th}	95^{th}
$1(\text{Diversify})_{g,n,t}$	13,018,734	0.00	0.07	0.00	0.00	0.00	0.00	0.00
$\text{HC Distance}_{g,n,t\text{-}1}$	13,018,734	0.93	0.10	0.74	0.92	0.97	0.99	1.00
$\text{HC Distance}_{g,n,t\text{-}1}$ (wages)	11,736,112	0.93	0.10	0.73	0.91	0.96	0.99	1.00
$\text{HC Distance}_{g,n,t\text{-}1}$ (hours)	13,008,638	0.93	0.10	0.74	0.92	0.97	0.99	1.00
$\text{HC Distance}_{g,n,t0}$	12,028,974	0.94	0.09	0.78	0.93	0.97	0.99	1.00
$\text{HC Distance}_{g,n,t1}$	13,018,734	0.82	0.15	0.52	0.76	0.87	0.93	0.98

Figure 2. Human Capital Distance Across Sectors

Source: Matched employer-employee dataset. *Sample:* Firms that enter a new sector and realize at least 5% of their sales in the new sector in the 2005-2014 period.

HEATMAP GOES HERE

Table 3. Human Capital Distance and Diversification Sector Selection

Source: SDC Platinum, BvD Zephyr, EAE/VAC survey, tax files, ownership links dataset, and matched employer-employee dataset. *Sample:* Firms that enter a new sector in the 2005-2014 period. The table reports OLS estimates and analyzes the effect of human capital on the type of diversification strategy. The dependent variable $\mathbb{1}(\text{Diversification})_{g,n,t}$ is a dummy variable that takes value one if firm g enters sector n in year t. Entries are identified with sales reported at the 3-digit level of the French classification of industries. The main independent variable is human capital distance, $HC\ Distance_{g,n,t-1}$. The variable measures the distance of the vector of occupations of firm g to the vector of occupations of sector n in year $t-1$. Control variables include the number of workers in logarithms as well as value added, number of occupations, total cash holdings, tangible assets, and total wages, all five scaled by the number of workers employed by the firm. Origin denotes the sector in which the firm realizes most of its sales at $t-1$. Entry is the sector that the firm enters in year t. Year is the date of entry. Standard errors are clustered at the sector of origin, sector of entry, and firm level and are reported in parentheses. *, **, and *** denote results that are significantly different from zero at the 10, 5 and 1% levels, respectively.

Dependent variable:	$\mathbb{1}(\text{Diversification})_{g,n,t}$			
	(1)	(2)	(3)	(4)
HC Distance$_{g,n,t-1}$	-0.023***	-0.058***	-0.059***	-0.029***
	(0.003)	(0.007)	(0.007)	(0.003)
Value added/N. Workers$_{g,t-1}$	-0.000	-0.000		
	(0.001)	(0.003)		
N. Occupations/N. Workers$_{g,t-1}$	-0.001***	-0.005***		
	(0.000)	(0.001)		
log(N. Workers)$_{g,t-1}$	-0.000***	-0.001***		
	(0.000)	(0.000)		
Tangible Assets/N. Workers$_{g,t-1}$	-0.000	-0.001		
	(0.000)	(0.001)		
Total wages/N. Workers$_{g,t-1}$	-0.002	0.002		
	(0.004)	(0.009)		
Cash/N. Workers$_{g,t-1}$	-0.001	-0.001		
	(0.001)	(0.003)		
Controls	Yes	Yes	No	No
Origin-Entry-Year FE	Yes	No	No	Yes
Firm FE	No	Yes	Yes	No
Firm-Year FE	No	No	No	Yes
R^2	0.125	0.003	0.002	0.120
Observations	13,018,734	13,056,293	13,515,402	13,480,046

Table 4. Human Capital Distance and Build or Buy

Source: SDC Platinum, BvD Zephyr, EAE/VAC survey, tax files, ownership links dataset, and matched employer-employee dataset. *Sample:* Firms that enter a new sector and realize at least 5% of their sales in the new sector in the 2005-2014 period. The table reports OLS estimates and analyzes the effect of human capital on the type of diversification strategy. The dependent variable $\mathbb{1}(\text{Buy})_{g,n,t}$ is a dummy variable that takes value one if the entry of firm g in sector n in year t is made by buying and 0 if it is made by building. Entries are identified with sales reported at the 3-digit level of the French classification of industries. The main independent variable is the human capital distance, *HC Distance*$_{g,n,t-1}$.The variable measures the distance of the vector of occupations of firm g to the vector of occupations of sector n in year $t-1$. Control variables include the number of workers in logarithms as well as value added, the number of occupations, total cash holdings, tangible assets, and total wages, all five scaled by the number of workers employed by the firm. Origin denotes the sector in which the firm realizes most of its sales at $t-1$. Origin (second) denotes the second sector (in terms of sales) in which the firm is present at $t-1$. Entry is the sector that the firm enters in year t. Year is the date of entry. Size is the quartile of sales in the main sector of origin. Standard errors are double-clustered at the sector of origin and sector of entry level and are reported in parentheses. *, **, and *** denote results that are significantly different from zero at the 10, 5 and 1% levels, respectively.

Dependent variable:	$\mathbb{1}(\text{Buy})_{g,n,t}$				
	(1)	(2)	(3)	(4)	(5)
HC Distance$_{g,n,t-1}$	0.018**	0.026***	0.022***	0.017**	0.010**
	(0.009)	(0.009)	(0.007)	(0.007)	(0.005)
N. Occupations/N. Employees$_{g,t-1}$		0.014**	0.016***	0.007**	-0.025**
		(0.006)	(0.006)	(0.003)	(0.011)
log(N. Employees)$_{g,t-1}$		0.008***	0.010***	0.005**	-0.049***
		(0.002)	(0.003)	(0.002)	(0.011)
Tangible Assets/N. Employees$_{g,t-1}$		-0.074***	-0.074***	-0.028**	-0.038
		(0.028)	(0.024)	(0.014)	(0.052)
Cash/N. Employees$_{g,t-1}$		0.109***	0.110**	0.032	-0.050
		(0.040)	(0.044)	(0.028)	(0.086)
Value added/N. Employees$_{g,t-1}$		-0.004	0.004	0.017	-0.232**
		(0.061)	(0.059)	(0.053)	(0.104)
Total wages/N. Employees$_{g,t-1}$		0.790**	0.828**	0.470**	0.471
		(0.374)	(0.352)	(0.195)	(0.475)
Controls	No	Yes	Yes	Yes	Yes
Origin × Entry × Year FE	Yes	Yes	No	No	No
Origin × Size × Entry FE	No	No	Yes	No	No
Origin (main) × Origin (second) × Entry FE	No	No	No	Yes	No
Origin × Entry FE	No	No	No	No	Yes
Firm FE	No	No	No	No	Yes
Year FE	No	No	Yes	Yes	Yes
Adjusted R^2	0.171	0.179	0.130	0.226	0.697
Observations	32,936	32,936	31,366	26,328	18,791

Table 5. Tests of Potential Sample Selection

Source: SDC Platinum, BvD Zephyr, EAE/VAC survey, tax files, ownership links dataset, and matched employer-employee dataset. *Sample:* Firms that enter a new sector in the 2005-2014 period. The table reports OLS estimates and analyzes the effect of human capital on the type of diversification strategy. The dependent variable $\mathbb{1}(\text{Buy})_{g,n,t}$ is a dummy variable that takes value one if the entry of firm g in sector n in year t is made by buying and 0 if it is made by building. Entries are identified with sales reported at the 3-digit level of the French classification of industries. The main independent variable is human capital distance, $HC\ Distance_{g,n,t-1}$. The variable measures the distance of the vector of occupations of firm g to the vector of occupations of sector n in year $t-1$. Control variables include the number of workers in logarithms as well as value added, number of occupations, total cash holdings, tangible assets, and total wages, all five scaled by the number of workers employed by the firm. Origin denotes the sector in which the firm realizes most of its sales at $t-1$. Entry is the sector that the firm enters in year t. Year is the date of entry. In Panel A, we replace the human capital variable using the wage and hour shares in the computation of the scalar product (columns 2 and 3) and by taking the value of the human capital distance for the first available observation of the firm in the sample (column 4). In Panel B, we instrument the human capital distance variable by the distance predicted by its sectoral repartition of sales. In Panel C, we test the robustness of our results to the size of entry into the new sector. Size of entry is measured as the ratio of sales realized in the new sector relative to total sales of firm g in year $t-1$. Standard errors are double-clustered at the sector of origin and sector of entry level and are reported in parentheses. *, **, and *** denote results that are significantly different from zero at the 10, 5 and 1% levels, respectively.

SHIFT ACTIVITY AND SERIAL ACQUIRER TABLE HERE

Table 6. Human Capital Distance: Tests of Endogeneity

Source: SDC Platinum, BvD Zephyr, EAE/VAC survey, tax files, ownership links dataset, and matched employer-employee dataset. *Sample:* Firms that enter a new sector in the 2005-2014 period. The table reports OLS estimates and analyzes the effect of human capital on the type of diversification strategy. The dependent variable $\mathbb{1}(\text{Buy})_{g,n,t}$ is a dummy variable that takes value one if the entry of firm g in sector n in year t is made by buying and 0 if it is made by building. Entries are identified with sales reported at the 3-digit level of the French classification of industries. The main independent variable is human capital distance, $HC\ Distance_{g,n,t-1}$. The variable measures the distance of the vector of occupations of firm g to the vector of occupations of sector n in year $t-1$. Control variables include the number of workers in logarithms as well as value added, number of occupations, total cash holdings, tangible assets, and total wages, all five scaled by the number of workers employed by the firm. Origin denotes the sector in which the firm realizes most of its sales at $t-1$. Entry is the sector that the firm enters in year t. Year is the date of entry. In Panel A, we replace the human capital variable using the wage and hour shares in the computation of the scalar product (columns 2 and 3) and by taking the value of the human capital distance for the first available observation of the firm in the sample (column 4). In Panel B, we instrument the human capital distance variable by the distance predicted by its sectoral repartition of sales. In Panel C, we test the robustness of our results to the size of entry into the new sector. Size of entry is measured as the ratio of sales realized in the new sector relative to total sales of firm g in year $t-1$. Standard errors are double-clustered at the sector of origin and sector of entry level and are reported in parentheses. *, **, and *** denote results that are significantly different from zero at the 10, 5 and 1% levels, respectively.

HC Distance ENDOGENEITY TABLE HERE

Table 7. Build or Buy: More Evidence on the Role of Labor

Source: SDC Platinum, BvD Zephyr, EAE/VAC survey, tax files, ownership links dataset, and matched employer-employee dataset. *Sample:* Firms that enter a new sector in the 2005-2014 period. The table reports OLS estimates and analyzes the effect of human capital on the type of diversification strategy. The dependent variable $\mathbb{1}(\text{Buy})_{g,n,t}$ is a dummy variable that takes value one if the entry of firm g in sector n in year t is made by buying and 0 if it is made by building. Entries are identified with sales reported at the 3-digit level of the French classification of industries. The main independent variable is human capital distance, $HC\ Distance_{g,n,t-1}$. The variable measures the distance of the vector of occupations of firm g to the vector of occupations of sector n in year $t-1$. Control variables include the number of workers in logarithms as well as value added, number of occupations, total cash holdings, tangible assets, and total wages, all five scaled by the number of workers employed by the firm. Origin denotes the sector in which the firm realizes most of its sales at $t-1$. Entry is the sector that the firm enters in year t. Year is the date of entry. In Panel A, we replace the human capital variable using the wage and hour shares in the computation of the scalar product (columns 2 and 3) and by taking the value of the human capital distance for the first available observation of the firm in the sample (column 4). In Panel B, we instrument the human capital distance variable by the distance predicted by its sectoral repartition of sales. In Panel C, we test the robustness of our results to the size of entry into the new sector. Size of entry is measured as the ratio of sales realized in the new sector relative to total sales of firm g in year $t-1$. Standard errors are double-clustered at the sector of origin and sector of entry level and are reported in parentheses. *, **, and *** denote results that are significantly different from zero at the 10, 5 and 1% levels, respectively.

LABOR MARKET TIGHTNESS INTERACTION + DELTA WORKERS AND HC DISTANCE BEFORE/AFTER. change tightness measure.

Table 8. Build or Buy: Cross-Sectional Variations

Source: SDC Platinum, BvD Zephyr, EAE/VAC survey, tax files, ownership links dataset, and matched employer-employee dataset. *Sample:* Firms that enter a new sector in the 2005-2014 period. The table reports OLS estimates and analyzes the effect of human capital on the type of diversification strategy. The dependent variable $\mathbb{1}(\text{Buy})_{g,n,t}$ is a dummy variable that takes value one if the entry of firm g in sector n in year t is made by buying and 0 if it is made by building. Entries are identified with sales reported at the 3-digit level of the French classification of industries. The main independent variable is human capital distance, $HC\ Distance_{g,n,t-1}$. The variable measures the distance of the vector of occupations of firm g to the vector of occupations of sector n in year $t-1$. Control variables include the number of workers in logarithms as well as value added, number of occupations, total cash holdings, tangible assets, and total wages, all five scaled by the number of workers employed by the firm. Origin denotes the sector in which the firm realizes most of its sales at $t-1$. Entry is the sector that the firm enters in year t. Year is the date of entry. In Panel A, we replace the human capital variable using the wage and hour shares in the computation of the scalar product (columns 2 and 3) and by taking the value of the human capital distance for the first available observation of the firm in the sample (column 4). In Panel B, we instrument the human capital distance variable by the distance predicted by its sectoral repartition of sales. In Panel C, we test the robustness of our results to the size of entry into the new sector. Size of entry is measured as the ratio of sales realized in the new sector relative to total sales of firm g in year $t-1$. Standard errors are double-clustered at the sector of origin and sector of entry level and are reported in parentheses. *, **, and *** denote results that are significantly different from zero at the 10, 5 and 1% levels, respectively.

CROSS-SECTIONAL TESTS TABLE HERE: New firm creations, Growth sector sales, Competition, Capital intensity

Table 9. Human Capital Distance and Post-Entry Performance

Source: SDC Platinum, BvD Zephyr, EAE/VAC survey, tax files, ownership links dataset, and matched employer-employee dataset. *Sample:* Firms that enter a new sector in the 2005-2014 period. The table reports OLS estimates and analyzes the effect of human capital on the type of diversification strategy. The dependent variable $\mathbb{1}(\text{Buy})_{g,n,t}$ is a dummy variable that takes value one if the entry of firm g in sector n in year t is made by buying and 0 if it is made by building. Entries are identified with sales reported at the 3-digit level of the French classification of industries. The main independent variable is human capital distance, $HC\ Distance_{g,n,t-1}$. The variable measures the distance of the vector of occupations of firm g to the vector of occupations of sector n in year $t-1$. Control variables include the number of workers in logarithms as well as value added, number of occupations, total cash holdings, tangible assets, and total wages, all five scaled by the number of workers employed by the firm. Origin denotes the sector in which the firm realizes most of its sales at $t-1$. Entry is the sector that the firm enters in year t. Year is the date of entry. In Panel A, we replace the human capital variable using the wage and hour shares in the computation of the scalar product (columns 2 and 3) and by taking the value of the human capital distance for the first available observation of the firm in the sample (column 4). In Panel B, we instrument the human capital distance variable by the distance predicted by its sectoral repartition of sales. In Panel C, we test the robustness of our results to the size of entry into the new sector. Size of entry is measured as the ratio of sales realized in the new sector relative to total sales of firm g in year $t-1$. Standard errors are double-clustered at the sector of origin and sector of entry level and are reported in parentheses. *, **, and *** denote results that are significantly different from zero at the 10, 5 and 1% levels, respectively.

TABLE POST-ENTRY PERFORMANCE HERE

Appendix A Theoretical Appendix

A.1 Main Prediction 1

Proof. For each sector i, the firm maximizes over its price subject to the demand function and the marginal cost of production:

$$\max_{p_i} \alpha_i q_i (p_i - c(i)), \tag{A1}$$

where q_i is given by (1), $c(i)$ is given by (2), and the value of α_i is one if the firm enters by building and α_E if it enters by buying. The first-order condition of this problem is

$$q_i^* = \frac{a - c(i)}{2b}. \tag{A2}$$

Substituting the q_i^* given by (A2) into the firm's profit in sector i given by (A1) and taking the integral over all sectors i, we add the fixed cost of production and the hiring costs to obtain the firm's total profit:

$$\frac{\alpha_k}{4b}(a - c(i))^2 - F - hm_i^2, \tag{A3}$$

where (A3) denotes the firm's total profit from building if $\alpha_k = 1$ and the firm's total profit from buying if $\alpha_k = \alpha_E$. The maximization of (A3) with respect to m_i yields the first-order condition[36]

$$m_i^* = \frac{\alpha_i/4b}{h - \alpha_i/4b} \left(a - (c_0 + \gamma_l i)\right). \tag{A4}$$

By rearranging (A4) and plugging it into (A3), we can rewrite the profit from entering a new sector i by building as in (3) and the profit from entering a new sector by buying an incumbent in sector i as in (4).

The firm faces two options for diversification. First, it can diversify in close sectors by buying and in distant sectors by building. Without loss of generality, we define a sector $z > 0$ and a value $\epsilon > 0$ such that the firm buys on $[z - \epsilon, z]$ and builds on $[z, z + \epsilon]$. We denote by

[36] We assume that $h > \frac{1}{4b}$ to ensure that the first-order condition yields a maximum. Otherwise, if $h < \frac{1}{4b}$, the first-order condition yields a minimum. In that case, the costs of hiring new employees are so low that the firm wants to hire infinite new employees and reduce its marginal cost of production to zero. In practice, we do not observe firms hiring infinite employees to enter a new sector. See Section 6.1 for a discussion of the search costs induced by labor scarcity.

$\pi_1 = \int_{z-\epsilon}^{z} \pi^{Buy}(i)di + \int_{z}^{z+\epsilon} \pi^{Build}(i)di$ the firm's profit under the first option.

Second, the firm can diversify in close sectors by building and in distant sectors by buying. In that case, the firm builds on $[z-\epsilon, z]$ and buys on $[z, z+\epsilon]$. We denote by $\pi_2 = \int_{z-\epsilon}^{z} \pi^{Build}(i)di + \int_{z}^{z+\epsilon} \pi^{Buy}(i)di$ the firm's profit under the second option.

We then compare the firm's profit under the first and second strategies as follows:

$$\pi_1 - \pi_2 = \frac{\alpha_E}{4b\gamma_E} \frac{h}{h - \frac{\alpha_E}{4b}} f(\gamma_E, z) - \frac{1}{4b\gamma_I} \frac{h}{h - \frac{1}{4b}} f(\gamma_I, z),$$

where we define $f(\gamma_i, z) \equiv [(g(\gamma_i, z+\epsilon) - g(\gamma_i, z)) - (g(\gamma_i, z) - g(\gamma_i, z-\epsilon))]$, with $g(\gamma_i, z) \equiv \frac{1}{3}\left(a - (c_0 + \gamma z)\right)^3$. We can rewrite $f(\gamma_i, z)$ as $f(\gamma_i, z) = 6\epsilon^2\gamma_i^2(a - c_0 - \gamma_i z)$. $f(\gamma_i, z)$ is increasing in γ_i as long as $z \leq \frac{2}{3\gamma_i}(a - c_0)$, which is the case whenever the fixed costs of entering a new sector, F, are large enough. Assuming that this condition is met and using $\gamma_I > \gamma_E$, it follows that $f(\gamma_I, z) > f(\gamma_E, z)$. Therefore,

$$\pi_1 - \pi_2 < f(\gamma_I, z) \frac{h}{4b\gamma_E} \left(\frac{\alpha_E}{h - \frac{\alpha_E}{4b}} - \frac{1}{h - \frac{1}{4b}} \right),$$

which is always negative because the share of profits that goes to the acquiring firm is $\alpha_E < 1$. As a result, the firm chooses strategy 2 and enters close sectors by building and distant sectors by buying. QED. □

A.2 Main Prediction 2

Proof. From Section A.1, we know that the firm will optimally diversify by building in sectors that are close to its core competence and by buying in sectors that are distant.

We denote by $[0, \beta]$ the set of close sectors the firm enters by building and by $\left[\beta, \delta^{Buy}\right]$ the set of distant sectors the firm enters by buying an incumbent company in sector i. By rearranging (A4) and plugging it into (A3), the buying firm's profits can be rewritten as follows:

$$\pi^{Buy}(i) = \int_0^{\beta} \frac{1}{4b}(a - (c_0 + \gamma_I i))^2 di + \int_{\beta}^{\delta^{Buy}} \frac{\alpha_E}{4b}(a - (c_0 + \gamma_E i))^2 di - \int_0^{\delta^{Buy}} F di, \qquad \text{(A5)}$$

where the first part of (A5) is the firm's profits from building in the set of close sectors $[0, \beta]$ and the second part is the profits from buying in the set of more distant sectors $\left[\beta, \delta^{Buy}\right]$. The third part of (A5) is the sum of fixed costs the firm must incur to enter new sectors.

The first-order conditions with respect to β and δ imply the equilibrium values

$$\beta^* = \left(\frac{1-K}{\gamma_I - K\gamma_E}\right)(a - c_0), \tag{A6}$$

$$\delta^{*,Buy} = \frac{a - c_0 - H(\gamma_E)}{\gamma_E}, \tag{A7}$$

where $K \equiv \frac{\alpha_E\left(h-\frac{1}{4b}\right)}{h-\frac{\alpha_E}{4b}} < 1$; thus, $\beta^* > 0$, and $H(x) \equiv \left(\frac{4bF\left(h-\frac{x}{4b}\right)}{xh}\right)^{1/2}$.

Using the equilibrium values of β^* and $\delta^{*,Buy}$ given by (A6) and (A7), we find that the firm only diversifies by building in close sectors *without* diversifying by buying in distant sectors whenever $\beta^* > \delta^{*,Buy}$, that is, if $c_0 \geq \underline{c_0} = a - \frac{H(\gamma_E)}{\gamma_E G}$, where $G \equiv \left(\frac{1}{\gamma_E} - \frac{1-K}{\gamma_I - K\gamma_E}\right) > 0$. The scope of the building-only firm is then

$$\delta^{*,Build} = \frac{a - c_0 - H(1)}{\gamma_I}. \tag{A8}$$

From (A8), it is straightforward to show the firm does not diversify and only produces in its core sector whenever $\delta^{*,Build} = 0$, i.e., if $c_0 \geq \overline{c_0} = a - \frac{H(1)}{\gamma_I}$. Main Prediction 2 obtains. QED. □

Appendix B Description of Variables

Variables	Description
Dependent variables	
$\mathbb{1}(\text{Buy})_{f,n,t}$	Dummy variable that takes value one if entry into sector n is made through an acquisition. *Source:* SDC Platinum, Zephyr, ownership links dataset, ESA survey.
HC Distance$_{g,n,t}$	Human capital distance is given by equation (5): $$\text{HC distance}_{g,n,t} = 1 - \frac{\sum_i s_{g,i,t} \cdot s_{n,i,t}}{\sqrt{\sum_i s_{g,i,t}^2}\sqrt{\sum_i s_{n,i,t}^2}},$$ where $s_{g,I,t}$ is the share of employees in firm g employed in occupation i, and $s_{n,I,t}$ is the share of employees in sector n employed in occupation i.
HC Distance$_{g,n,t}$ (wages)	Human capital distance as given by (5), where $s_{g,I,t}$ is the share of firm g's total wage bill that goes to employees in occupation i, and $s_{n,I,t}$ is the share of sector n's total wage bill that goes to employees in occupation i.
HC Distance$_{g,n,t}$ (hours)	Human capital distance as given by (5), where $s_{g,I,t}$ is the share of the total number of hours worked in firm g by employees in occupation i, and $s_{n,I,t}$ is the share of the total number of hours worked in sector n by employees in occupation i.
HC Distance$_{g,n,t_0}$	HC Distance value the first year firm g appears in the sample, such that $$\textit{HC distance}_{g,n,t_0} = \sum_i s_{g,i,t_0} \cdot s_{n,i,t_0}.$$
HC Distance$^{Sector}_{o,n,t}$	Weighted sum of the HC distance between sectors o and sector n, weighted by firm g's share of sales realized in sector o: $$\textit{HC distance}^{Sector}_{o,n,t} = \sum_o w_{o,g,t} \cdot \left(1 - \sum_i s_{o,i,t} \cdot s_{n,i,t}\right)$$, where $w_{o,g,t}$ is the share of sales realized in sector o by firm g in year t, and $s_{o,i,t}$ and $s_{n,i,t}$ are the shares of employees in sectors o and n, respectively, employed in occupation i.
Value added/N.workers$_{g,t-1}$	Total value added generated by firm g in year $t-1$, scaled by the number of workers in the firm. *Source:* Tax files, Ownership links dataset, Matched employer-employee dataset.
N. Occupations/N. Workers$_{g,t-1}$	Number of occupations in firm g in year $t-1$, scaled by the number of workers in the firm. *Source:* Tax files, Ownership links dataset, Matched employer-employee dataset.
log(N. Workers)$_{g,t-1}$	Logarithm of the number of workers in firm g in year $t-1$. *Source:* Matched employer-employee dataset, Ownership links dataset.
Tangible Assets/N. Workers$_{g,t-1}$	Total value of tangible assets held by firm g in year $t-1$, scaled by the number of workers in the firm. *Source:* Tax files, Ownership links dataset, Matched employer-employee dataset.
Total wages/N. Workers$_{g,t-1}$	Total wages of firm g in year $t-1$, scaled by the number of workers in the firm. *Source:* Tax files, Ownership links dataset, Matched employer-employee dataset.
Cash/N. Workers$_{g,t-1}$	Total cash holdings of firm g in year $t-1$, scaled by the number of workers in the firm. *Source:* Tax files, Ownership links dataset, Matched employer-employee dataset.
Entry share new sector$_{g,n,t-1}$	?
$\mathbb{1}(\text{Diversified})_{g,t-1}$	Dummy variable equal to one if firm g is diversified in year $t-1$, i.e., if it is operating in more than one sector where a sector is defined at the 3-digit level.
$\mathbb{1}(\text{Public})_{g,t-1}$	Dummy variable equal to one if at least one entity within the firm is a publicly listed company, zero otherwise. *Source:* Bureau van Dijk Amadeus
$\mathbb{1}(\text{Vertical link} \geq 1\%)_{o,n}$	Dummy variable that is equal to one if the sector of origin of the firm (as measured in the I-O OECD classification) sources more than 1% of its inputs from the sector of entry (as measured in the I-O OECD classification). The variable is not defined when the sector of origin and the sector of entry belong to the same item of the I-O OECD classification. The method follows Fan and Goyal's (2006) measure of vertical relatedness. *Source:* Ownership links dataset, Tax files, OECD 1999 French Input-Output table.

Continued next page

Description of Variables (continued)

Variables	Description
$\mathbb{1}(\text{Vertical link} \geq 5\%)_{o,n}$	Dummy variable that is equal to one if the sector of origin of the firm (as measured in the I-O OECD classification) sources more than 5% of its inputs from the sector of entry (as measured in the I-O OECD classification). The variable is not defined when the sector of origin and the sector of entry belong to the same item of the I-O OECD classification. The method follows Fan and Goyal's (2006) measure of vertical relatedness. *Source:* Ownership links dataset, Tax files, OECD 1999 French Input-Output table.
$\mathbb{1}(\text{New department})_{g,t}$ and $\mathbb{1}(\text{New region})_{g,t}$	Dummy variable that indicates whether firm g enters a new geographic zone in year t – whether a new department or a new region. France is divided into 101 departments and 25 regions over the sample period. A firm enters a new geographic zone if (i) it creates a subsidiary in year t that is located in a new region, (ii) it creates a plant in a new region in year t, or (iii) it acquires an existing firm in year t that is operating in a geographic zone in which the firm was not present at $t-1$. Source: Matched employer-employee dataset, Ownership links dataset, SIRENE.
$\text{Labor scarcity}_{g,n,t-1}$	Sales-weighted average exposure to the sector–commuting zone labor scarcity: $$\textit{Labor scarcity}_{g,n,t} = \sum_z \omega_{z,g,t} \cdot \sum_i \textit{Scarce occupation}_{i,z,t} \cdot s_{n,i,t}$$, where $\omega_{z,g,t-1}$ is the share of sales realized by firm g in zone z in year t, $\textit{Scarce occupation}_{i,z,t}$ is a dummy variable that identifies a scarce occupation i in commuting zone z in year t, and $s_{n,i,t}$ is the share of employees in sector n employed in occupation i.
$\text{N. New firms}_{n,t-1}$	Number of firms newly created in the sector of entry in year $t-1$. This measure is used as a proxy for growth opportunities.

9. Additional Tables for Internet Appendix

Table A1. Robustness Checks: Alternative Measures of Human Capital Distance

Source: SDC Platinum, BvD Zephyr, EAE/VAC survey, tax files, ownership links dataset, and matched employer-employee dataset. *Sample:* Firms that enter a new sector in the 2005-2014 period. The table reports OLS estimates and analyzes the effect of human capital on the type of diversification strategy. The dependent variable $\mathbb{1}(\text{Buy})_{g,n,t}$ is a dummy variable that takes value one if the entry of firm g in sector n in year t is made by buying and 0 if it is made by building. Entries are identified with sales reported at the 3-digit level of the French classification of industries. The main independent variable is human capital distance, *HC Distance*$_{g,n,t-1}$. The variable measures the distance of the vector of occupations of firm g to the vector of occupations of sector n in year $t-1$. Control variables include the number of workers in logarithms as well as value added, number of occupations, total cash holdings, tangible assets, and total wages, all five scaled by the number of workers employed by the firm. Origin denotes the sector in which the firm realizes most of its sales at $t-1$. Entry is the sector that the firm enters in year t. Year is the date of entry. In Panel A, we replace the human capital variable using the wage and hour shares in the computation of the scalar product (columns 2 and 3) and by taking the value of the human capital distance for the first available observation of the firm in the sample (column 4). In Panel B, we instrument the human capital distance variable by the distance predicted by its sectoral repartition of sales. In Panel C, we test the robustness of our results to the size of entry into the new sector. Size of entry is measured as the ratio of sales realized in the new sector relative to total sales of firm g in year $t-1$. Standard errors are double-clustered at the sector of origin and sector of entry level and are reported in parentheses. *, **, and *** denote results that are significantly different from zero at the 10, 5 and 1% levels, respectively.

TABLE WITH ROBUSTNESS OF HC MEASURE HERE: wages, hours worked, other similarities?

Table A2. Robustness Checks: Alternative Entry Thresholds for Build Entries

Source: SDC Platinum, BvD Zephyr, EAE/VAC survey, tax files, ownership links dataset, and matched employer-employee dataset. *Sample:* Firms that enter a new sector in the 2005-2014 period. The table reports OLS estimates and analyzes the effect of human capital on the type of diversification strategy. The dependent variable $\mathbb{1}(\text{Buy})_{g,n,t}$ is a dummy variable that takes value one if the entry of firm g in sector n in year t is made by buying and 0 if it is made by building. Entries are identified with sales reported at the 3-digit level of the French classification of industries. The main independent variable is human capital distance, *HC Distance*$_{g,n,t-1}$. The variable measures the distance of the vector of occupations of firm g to the vector of occupations of sector n in year $t-1$. Control variables include the number of workers in logarithms as well as value added, number of occupations, total cash holdings, tangible assets, and total wages, all five scaled by the number of workers employed by the firm. Origin denotes the sector in which the firm realizes most of its sales at $t-1$. Entry is the sector that the firm enters in year t. Year is the date of entry. In Panel A, we replace the human capital variable using the wage and hour shares in the computation of the scalar product (columns 2 and 3) and by taking the value of the human capital distance for the first available observation of the firm in the sample (column 4). In Panel B, we instrument the human capital distance variable by the distance predicted by its sectoral repartition of sales. In Panel C, we test the robustness of our results to the size of entry into the new sector. Size of entry is measured as the ratio of sales realized in the new sector relative to total sales of firm g in year $t-1$. Standard errors are double-clustered at the sector of origin and sector of entry level and are reported in parentheses. *, **, and *** denote results that are significantly different from zero at the 10, 5 and 1% levels, respectively.

TABLE WITH ROBUSTNESS OF ENTRY THRESHOLD – former Table 4 panel C

Table A3. The Role of Product Market and Geographic Distances

Source: SDC Platinum, BvD Zephyr, EAE/VAC survey, tax files, ownership links dataset, and matched employer-employee dataset. *Sample:* Firms that enter a new sector in the 2005-2014 period. The table reports OLS estimates and analyzes the effect of human capital on the type of diversification strategy. The dependent variable $\mathbb{1}(\text{Buy})_{g,n,t}$ is a dummy variable that takes value one if the entry of firm g in sector n in year t is made by buying and 0 if it is made by building. Entries are identified with sales reported at the 3-digit level of the French classification of industries. The main independent variable is human capital distance, *HC Distance*$_{g,n,t-1}$. The variable measures the distance of the vector of occupations of firm g to the vector of occupations of sector n in year $t-1$. Control variables include the number of workers in logarithms as well as value added, number of occupations, total cash holdings, tangible assets, and total wages, all five scaled by the number of workers employed by the firm. Origin denotes the sector in which the firm realizes most of its sales at $t-1$. Entry is the sector that the firm enters in year t. Year is the date of entry. In columns 1 and 2, we interact the human capital distance with an indicator equal to one if the sector of origin is in the same 1- or 2-digit category of the French classification of sectors. Columns 3 and 4 include Fan and Goyal's (2006) measure of vertical relatedness. It measures the intensity of vertical links between the main sector of activity of firm g in year $t-1$ and the sector that g enters in year t. "Vertical" is a dummy variable that takes value 1 if the vertical relatedness exceeds 5% or 10%. Columns 5 and 6 include a dummy that indicates whether the firm enters a new geographic zone in year t. We use "department" and "region" as definitions of the geographic zone (France is divided into 25 regions and 96 departments over the sample period). A firm enters a new geographic zone if (i) the entering subsidiary is created in year t and located in a new region, (ii) the entering subsidiary opens a plant in a new region in year t, or (iii) the entering subsidiary is acquired in year t and operates in a region in which the firm was not present at $t-1$. Standard errors are double-clustered at the sector of origin and sector of entry level and are reported in parentheses. *, **, and *** denote results that are significantly different from zero at the 10, 5 and 1% levels, respectively.

Dependent variable:	$\mathbb{1}(\text{Buy})_{g,n,t}$					
	Product market distance		Vertical link		geographic distance	
	1-digit	2-digits	$> 5\%$	$> 10\%$	Department	Region
	(1)	(2)	(3)	(4)	(5)	(6)
HC Distance$_{g,n,t-1}$	0.040***	0.029***	0.034***	0.034***	0.022***	0.022***
	(0.012)	(0.009)	(0.012)	(0.011)	(0.007)	(0.007)
Same Industry$_{o,n}$	0.032***					
	(0.011)					
Same Ind.$_{o,n}$× HC Distance$_{g,n,t-1}$	-0.034***					
	(0.012)					
Same Industry$_{o,n}$		0.021**				
		(0.010)				
Same Ind.$_{o,n}$× HC Distance$_{g,n,t-1}$		-0.018				
		(0.012)				
Vertical link$_{o,n}$			0.025			
			(0.016)			
Link$_{o,n}$× HC Distance$_{g,n,t-1}$			-0.029*			
			(0.017)			
Vertical link$_{o,n}$				0.040**		
				(0.016)		
Link$_{o,n}$× HC Distance$_{g,n,t-1}$				-0.039**		
				(0.018)		
New area$_{g,t}$					0.130***	
					(0.040)	
New area$_{g,t}$× HC Distance$_{g,n,t-1}$					0.005	
					(0.043)	
New area$_{g,t}$						0.183***
						(0.056)
New area$_{g,t}$× HC Distance$_{g,n,t-1}$						0.009
						(0.066)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Origin × Year FE	Yes	Yes	Yes	Yes	No	No
Entry × Year FE	Yes	Yes	Yes	Yes	No	No
Origin × Entry × Year FE	No	No	No	No	Yes	Yes
Adjusted R^2	0.092	0.091	0.099	0.099	0.235	0.242
Observations	32,936	32,936	26,807	26,807	32,807	32,807

Table A4. Human Capital or Physical Capital?

Source: SDC Platinum, BvD Zephyr, EAE/VAC survey, tax files, ownership links dataset, and matched employer-employee dataset. *Sample:* Firms that enter a new sector in the 2005-2014 period. The table reports OLS estimates and analyzes the effect of human capital on the type of diversification strategy. The dependent variable $\mathbb{1}(\text{Buy})_{g,n,t}$ is a dummy variable that takes value one if the entry of firm g in sector n in year t is made by buying and 0 if it is made by building. Entries are identified with sales reported at the 3-digit level of the French classification of industries. The main independent variable is human capital distance, $HC\ Distance_{g,n,t-1}$. The variable measures the distance of the vector of occupations of firm g to the vector of occupations of sector n in year $t-1$. Control variables include the number of workers in logarithms as well as value added, number of occupations, total cash holdings, tangible assets, and total wages, all five scaled by the number of workers employed by the firm. Origin denotes the sector in which the firm realizes most of its sales at $t-1$. Entry is the sector that the firm enters in year t. Year is the date of entry. In Panel A, we replace the human capital variable using the wage and hour shares in the computation of the scalar product (columns 2 and 3) and by taking the value of the human capital distance for the first available observation of the firm in the sample (column 4). In Panel B, we instrument the human capital distance variable by the distance predicted by its sectoral repartition of sales. In Panel C, we test the robustness of our results to the size of entry into the new sector. Size of entry is measured as the ratio of sales realized in the new sector relative to total sales of firm g in year $t-1$. Standard errors are double-clustered at the sector of origin and sector of entry level and are reported in parentheses. *, **, and *** denote results that are significantly different from zero at the 10, 5 and 1% levels, respectively.

TABLE WITH PHYSICAL CAPITAL – Capital intensity sector, investment buckets

Table A5. Financing Constraints

Dependent variable:	$\mathbb{1}(\text{Buy})_{g,n,t}$			
Measure:	N.employees	Cash	Diversified group	Public group
	(1)	(2)	(3)	(4)
HC Distance$_{g,n,t\text{-}1}$	0.040***	0.035***	0.036***	0.029***
	(0.014)	(0.009)	(0.011)	(0.008)
2nd tercile of N. Workers$_{g,t\text{-}1}$	0.015**			
	(0.006)			
3rd tercile of N. Workers$_{g,t\text{-}1}$	0.031***			
	(0.008)			
2nd t. N. Workers$_{g,t\text{-}1}$× HC Distance$_{g,n,t\text{-}1}$	-0.025**			
	(0.010)			
3rd t. N. Workers$_{g,t\text{-}1}$× HC Distance$_{g,n,t\text{-}1}$	-0.026*			
	(0.013)			
2nd tercile of Cash/N. Workers$_{g,t\text{-}1}$		0.011*		
		(0.006)		
3rd tercile of Cash/N. Workers$_{g,t\text{-}1}$		0.015*		
		(0.009)		
2nd t. Cash/N. Workers$_{g,t\text{-}1}$× HC Distance$_{g,n,t\text{-}1}$		-0.011		
		(0.008)		
3rd t. Cash/N. Workers$_{g,t\text{-}1}$× HC Distance$_{g,n,t\text{-}1}$		-0.013		
		(0.009)		
1(Diversified)$_{g,t\text{-}1}$			0.018**	
			(0.008)	
1(Diversified)$_{g,t\text{-}1}$× HC Distance$_{g,n,t\text{-}1}$			-0.019**	
			(0.008)	
1(Public)$_{g,t\text{-}1}$				0.147**
				(0.061)
1(Public)$_{g,t\text{-}1}$× HC Distance$_{g,n,t\text{-}1}$				-0.084
				(0.081)
Controls	Yes	Yes	Yes	Yes
Origin × Entry × Year FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.118	0.118	0.119	0.125
Observations	32,936	32,936	32,936	32,936