Received: 8 June 2018 | Revised: 17 April 2020 | Accepted: 20 April 2020

DOI: 10.1002/smj.3200

RESEARCH ARTICLE

STRATEGIC MANAGEMENT JOURNAL WILEY

Internal agglomeration and productivity: Evidence from microdata

Evan Rawley[1] | **Robert Seamans**[2]

[1]Carlson School of Management, University of Minnesota, Minneapolis, Minnesota, USA

[2]Stern School of Business, New York University, New York, New York, USA

Correspondence
Robert Seamans, Stern School of Business, New York University, New York, NY, USA.
Email: rseamans@stern.nyu.edu

Funding information
Ewing Marion Kauffman Foundation

Abstract

Research Summary: We study how internal agglomeration—geographic clustering of business establishments owned by the same parent company—influences establishment productivity. Using Census microdata on the population of U.S. hotels from 1987-2007, we find that doubling the intensity of internal agglomeration is associated with a productivity increase of about 2% in pre-existing establishments. We consider several mechanisms that may be driving the productivity effect and find evidence consistent with the idea that an economically meaningful component of the productivity effect is due to knowledge transfer between internally agglomerated establishments. We replicate our main findings with Census microdata on the full population of U.S. restaurants from 1987-2007, suggesting that the internal agglomeration effects we document may generalize broadly to other industries with multi-unit firms.

Managerial Summary: Internal agglomeration is the geographic clustering of business establishments owned by the same parent company. This paper uses detailed Census data on hotels and restaurants to show how internal agglomeration influences performance. Interestingly, knowledge sharing between owned establishments in the same metropolitan area appears to be a key driver of the internal agglomeration effect.

KEYWORDS

agglomeration, corporate strategy, knowledge spillovers, productivity

1 | INTRODUCTION

Location decisions are crucial for multi-unit firms. One key element of location choice is whether to agglomerate internally—cluster business establishments owned by the same parent firm in close geographical proximity. While previous scholarship has suggested that internal agglomeration can be important (Alcacer & Delgado, 2016; Darr, Argote, & Epple, 1995), no large sample studies to date have linked internal agglomeration to establishment-level productivity effects, even though establishment productivity is known to be a particularly important determinant of firm performance (Syverson, 2011).

To address this lacuna in the literature, this paper uses establishment-level data from the Economic Census to evaluate whether there is large sample evidence of an internal agglomeration and productivity effect; and explores, conceptually and empirically, several potential mechanisms behind the hypothesized effect. The results show that internal agglomeration is associated with a 2% increase in productivity. Alternative specifications and increasingly demanding econometric tests, including matching on observables and instrumental variables techniques, all find a similar relationship between internal agglomeration and productivity. Taken together, a series of extensions and robustness checks suggest that an economically meaningful component of the productivity effect is due to knowledge transfer between internally agglomerated establishments, though we cannot precisely measure the relative contribution of all the potential mechanisms at play. We replicate our main findings on the U.S. food service ("restaurant") industry 1987–2007, suggesting that our findings generalize to other industries with multi-unit firms.

The paper makes two key contributions. First, we document a large sample, empirically rigorous relationship between internal agglomeration and productivity in two important multi-establishment industries. Second, we test a number of potential mechanisms behind the internal agglomeration effect and provide some evidence of intrafirm learning amongst internally agglomerated establishments. While much of the agglomeration literature focuses on the contemporaneous "flow" benefits of agglomeration—those benefits that accrue to the focal firm so long as establishments are agglomerated—such as complementary demand (Canina, Enz, & Harrison, 2005; Fabrizio & Thomas, 2012; McCann & Vroom, 2010) and resource sharing (Feldman, 2000; Folta, Cooper, & Baik, 2006; Krugman, 1991), this paper provides evidence that persistent "stock" benefits associated with internal agglomeration—those benefits that accrue to the focal firm, and remain in effect, even if internal agglomeration ceases—are also economically meaningful.[1]

Our analyses are compelling for several reasons. The data are unusually broad and rich—we have establishment-level data on a complete census of every hotel ever operated (for at least 5 years) in the United States 1987–2007. We use the data to offer relatively precise estimates of the magnitude and persistence of internal agglomeration. Moreover, by exploiting features of the empirical setting in our research design—particularly the fact that we study hundreds of distinct localized markets—we offer results that are robust to a wide range of alternative explanations. Furthermore, we triangulate our quantitative findings with qualitative insights gleaned from interviews with industry experts, which offer supporting evidence about the nature of the mechanisms behind the internal agglomeration effects we study.

[1]"Contemporaneous" versus "persistent" agglomeration benefits differ conceptually from "static" versus "dynamic" agglomeration benefits in that the former are defined with respect to a state of being—internally agglomerated or not—whereas the latter are defined with respect to time (Rosenthal & Strange, 2004).

2 | THEORY AND LITERATURE

2.1 | Internal and external agglomeration

The costs and benefits of agglomeration have fascinated scholars at least since Marshall (1920/1890) studied the relationship between firm performance and the geographical clustering of similar establishments. And, recent scholarship has convincingly linked external agglomeration to performance. For example, Greenstone, Hornbeck, and Moretti (2010) use quasi-exogenous entry of large manufacturing plants to show that incumbent firms in the same counties subsequently experience an increase in total factor productivity; and Henderson (2003) and Martin, Mayer, and Mayneris (2011) also find evidence of the link between external agglomeration and productivity. An excellent review by McCann and Folta (2008) provides many more examples.

While most research on agglomeration builds upon Marshall's work by focusing on interfirm agglomeration effects, agglomeration effects may be manifest within firms as well (Alcacer & Delgado, 2016; Darr et al., 1995). Conceptually, internal agglomeration effects roughly parallel interfirm agglomeration effects, except that internal effects are managed by the "visible hand" of the firm, while interfirm effects are essentially pure market externalities. Alcacer and Delgado (2016) study drivers of location choice for firms and argue that, in ignoring the important role of internal agglomeration, the prior literature has overestimated the effects of external agglomeration. Thus, the overarching objective of our paper is to systematically explore the understudied impact of internal agglomeration on firm performance, conceptually and empirically. In doing so, our paper complements Alcacer and Delgado (2016) and prior studies by providing a link between internal agglomeration and performance .

2.2 | Agglomeration and productivity mechanisms

Scholars have described a variety of mechanisms through which agglomeration leads to improved productivity. At a broad level, the key mechanisms include demand and supply responses, which may include monitoring effects, learning/knowledge spillovers, and resource sharing.[2] While we will describe context-specific examples of each potential internal agglomeration and productivity mechanism below, and provide tests that discriminate between features of each mechanism, the empirical analyses we provide will only be suggestive of which mechanisms are most important in the hotel industry.

2.2.1 | Demand responses

Agglomeration can benefit establishments by driving complementary demand (Canina et al., 2005; Chung & Kalnins, 2001; Fabrizio & Thomas, 2012; McCann & Vroom, 2010). For example, when enough restaurants operate in a neighborhood, that neighborhood may become known for dining, leading to increased customer traffic into the neighborhood to the benefit of all restaurants in the market (Glaeser, Kim, & Luca, 2017). It is reasonable to expect that internal agglomeration could deliver similar demand-side benefits too.

[2]See McCann and Vroom (2008) for a comprehensive review of the existing literature on these mechanisms.

2.2.2 | Supply responses

Agglomeration can also benefit establishments through supply-side responses. When more establishments enter an area, it becomes more likely that suppliers will gravitate to that area to serve the establishments (Alcacer & Chung, 2014; Folta et al., 2006; McCann & Folta, 2008). Greenstone et al. (2010) posit that local labor supply increases are likely one of the mechanisms underlying the link between external agglomeration and productivity increases in their empirical context, and it seems possible that internal agglomeration could lead to similar labor supply effects. For example, it may be the case that a new sister establishment needs to hire multiple workers in specialized roles. Some of these positions are filled by workers who respond to this demand by moving from outside the market area. An incumbent establishment could benefit from this increase in labor supply, particularly if they eventually employ the incoming labor. Moreover, even if the new establishment later exits the market, incumbents may continue to benefit from the increased labor supply so long as it does not quickly disappear following the exit of the new establishment.

2.2.3 | Monitoring

Increasing physical and informational frictions between establishments and headquarters tends to hamper productivity. For example, Henderson and Ono (2008) and Kalnins and Lafontaine (2013) show that increasing distance between an establishment and its headquarters increases monitoring costs, and Giroud (2013) demonstrates how easing travel between headquarters and a production facility leads to greater productivity. Intuitively, a distant corporate headquarters should be able to more easily monitor internally agglomerated establishments, compared to geographically dispersed establishments, by using the performance of each internally agglomerated establishment as a benchmark for the other. Thus, improved monitoring may be another mechanism through which internal agglomeration influences productivity.

2.2.4 | Learning/knowledge spillovers

The innovation literature provides many reasons why firms may benefit from agglomeration (Alcacer & Delgado, 2016; Feldman, 2000; Gilbert, McDougall, & Audretsch, 2008; Shaver & Flyer, 2000); for example, skilled labor may move across firm boundaries, taking productivity enhancing ideas from one firm to another (Feldman, 2000). Analogously, there are several reasons to expect that internal agglomeration may result in knowledge spillovers within the firm; for example, internal agglomeration may lead to sharing of best practices between geographically proximate establishments.

While it seems relatively straightforward to claim that new establishments probably learn from pre-existing establishments, it may also be the case that the relationship goes the other way: pre-existing establishments may learn from new sister establishments' as well. One empirically testable implication of geographically proximate internal learning is that if a pre-existing establishment learns from a sister establishment, the pre-existing establishment will continue to benefit from such knowledge, even if the sister establishment subsequently exits the market. We refer to such benefits as "persistent" effects *with respect to state* (i.e., being internally

agglomerated or not), to distinguish them from the contemporaneous "flow" benefits of internal agglomeration.

Of course, while internal and external agglomeration both can lead to increased learning between establishments in the same market, the strategic response of "treated" firms are different in each case; one might expect a firm to actively encourage knowledge spillovers across its internally agglomerated units, but closely guard knowledge from spilling over to geographically proximate units in other firms.

2.2.5 | Resource sharing

When a sister establishment is co-located with an incumbent, there are a variety of ways in which resources can be shared between them (Woo, Cannella Jr, & Mesquita, 2019). For example, local investments in facilities used to support both establishments, or intangible assets, such as marketing, advertising and other brand-building activities, can be shared between establishments. In general, therefore, one might expect that internal agglomeration will lead to productivity increases via scale and scope efficiencies from sharing common resources (Sakhartov, 2017).

2.3 | Contemporaneous versus persistent agglomeration effects

Considering the type of agglomeration effects (i.e., interfirm vs. internal) together with an establishment's agglomeration state (i.e., contemporaneous or not), suggests that there are four broad categories of agglomeration effects—interfirm contemporaneous effects, interfirm persistent effects, internal contemporaneous effects, and internal persistent effects (see Figure 1).

Interfirm contemporaneous agglomeration effects (quadrant I in Figure 1) include positive spillovers such as complementary demand whereby increased clustering drives customer traffic (McCann & Vroom, 2010), and complementary supply whereby increased clustering attracts specialized suppliers of resources (Saxenian, 1994). There may also be interfirm contemporaneous negative spillovers from increased competition for customers or resources (Alcacer, 2006), leading either to lower prices, as in standard Bertrand competition, or increased costs due to a reduction in bargaining power with suppliers (Brandenburger & Stuart, 2007).

Type/(state)	**Inter-firm (external) agglomeration**	**Internal agglomeration**
Contemporaneous (currently agglomerated)	**I** Benefits: Classic Marshallian agglomeration externalities, including complementary supply and demand Costs: Competition	**II** Benefits: Local economies of scale in procurement, reputation building, enhanced monitoring, ongoing learning Costs: Cannibalization
Persistent (agglomerated in the past, but not concurrently agglomerated)	**III** Benefits: Knowledge spillovers leaked from one firm to another Costs: Costly investments to imitate or deter rivals	**IV** Benefits: Knowledge spillovers shared between business units owned by the same parent firm Costs: Internal politicking, bureaucracy

FIGURE 1 Four types of agglomeration effects

Internal contemporaneous effects (quadrant II) include benefits that accrue to sister establishments owned by the same parent firm in close proximity to each other through economies of scale in purchasing and branding (Jin & Leslie, 2009), enhanced monitoring (Lu & Wedig, 2013), and ongoing learning and benchmarking (Baum & Ingram, 1998). There may also be internal contemporaneous negative effects, for example due to cannibalization of the customer base (Pancras, Sriram, & Kumar, 2012).

Persistent interfirm effects (quadrant III) are those external agglomeration benefits that are preserved even as agglomeration intensity declines. One class of persistent effects is due to knowledge spillovers. For example, leakage of a process or production technique or technological capability (Appleyard, 1996; Gilbert et al., 2008; Mowery, Oxley, & Silverman, 1996), and other long-lived changes, such as improved organizational practices or improved market position, but in principle could involve human as well as intellectual capital (Allen, 1984). Costs associated with persistent interfirm agglomeration may be characterized as adjustment costs and/or frictions. Adjustment costs could be physical, the costs associated with changing an establishment's location, and/or procedural, the costs associated with changing business practices, in response to the change in the intensity of agglomeration. More broadly, the costs might also include frictions associated with operating a business in the same way even as the intensity of agglomeration decreases—or the costs of failing to adjust.

Internal persistent effects (quadrant IV) are those internal agglomeration benefits that are preserved even when internal agglomeration ceases. Our interest in evaluating the evidence in favor of localized knowledge transfers, as opposed to knowledge across all business establishments, regardless of location, differentiates this work from the prior literature on within-firm learning, such as studies of the ways in which multinational corporations transmit knowledge from business establishments in one country to another (Bresman, Birkinshaw, & Nobel, 1999; Gupta & Govindarajan, 2000). Potential internal agglomeration costs include coordination, bureaucratization and internal politicking, costs that might also be persistent with respect to state (Rawley, 2010; Williamson, 2005; Zhou, 2011).

Figure 1 provides a roadmap of the main goals of the paper. First, we aim to measure the productivity effect of internal agglomeration, that is, the net (of costs and benefits) effect of the sum of contemporaneous and persistent effects (i.e., the sum of quadrants II and IV). While other papers have also measured the effect of internal agglomeration on various measures of performance (Darr et al., 1995), we believe our dataset provides an opportunity to develop a relatively precise estimate of the magnitude of internal agglomeration effects in a large, and economically meaningful, sample of establishments. Second, we aim to identify the relative contribution of contemporaneous (quadrant II) versus persistent (quadrant IV) effects of internal agglomeration. To the best of our knowledge, this is a novel contribution to the agglomeration literature. Our third goal is to explore which potential mechanisms may be driving the link between internal agglomeration and productivity. We do so by focusing on the mechanisms described above: supply and demand responses, monitoring effects, learning/knowledge spillovers, and resource sharing.

While we are keen to examine the mechanisms behind the proposed internal agglomeration effect, our first priority is simply to establish that internal agglomeration effects exist, and are economically meaningful. Therefore, our main hypothesis is:

Hypothesis 1 (H1). Internal agglomeration—clustering business establishments owned by the same parent firm in close geographical proximity—leads to improved productivity.

Below we evaluate the evidence for, or against, Hypothesis 1, using increasingly strict econometric tests, before exploring the potential mechanisms behind the hypothesis.

3 | DATA AND EMPIRICAL STRATEGY

3.1 | Data and setting

Our main empirical analyses center on the hotel industry—one of the largest and most visible industries with geographically clustered multi-unit firms. We then replicate our main findings in the restaurant industry, an industry with similar characteristics, as we describe in more detail in the discussion section. The data come from the 1987 to 2007 quinquennial (1987, 1992, 1997, 2002, and 2007) Economic Census of Services and Economic Census of Retail Trade, and include the revenue and payroll of every hotel and restaurant in the United States, tracked longitudinally with a unique identifier, along with location and firm identifiers.[3] Most establishments in the Economic Census are surveyed directly by the Census Bureau, though data are imputed for some establishments with few employees. We exclude establishments for which data has been imputed—establishments that are in the smallest 10% of the size distribution (by revenue)—along with other unusual observations—those in the 1st and 99th percentile of the labor factor productivity distribution.[4]

Before conducting large sample empirical tests, we conducted a number of exploratory interviews with industry executives, analysts, consultants and managers to better understand how internal agglomeration effects were manifest in the hotel industry. We found strong anecdotal support for previously documented benefits associated with internal agglomeration, including joint procurement, and more efficient monitoring. Interestingly, knowledge-based benefits were also noted.

For example, the chief financial officer of a large real estate investment trust (REIT) described how a process improvement in maid service—a reorganization of cleaning carts to speed room changeovers—spread across internally agglomerated establishments in one city long before other hotels in the same firm in other cities recognized the value of the change. Relatedly, a hotel manager with experience running a number of hotels in the Midwest discussed how standard operating procedures were often modified by hotel managers to account for market-level heterogeneity in customer preferences (e.g., with respect to check-in procedures and in-room amenities).

The idea that localized knowledge might pass back and forth between new and pre-existing establishments was buttressed by the director of business development at a prominent hospitality firm who explained that new establishments are often managed by seasoned managers from pre-existing hotels in the same market with deep connections to personnel at their pre-existing hotel. Because some customization is expected at a new establishment, it is common for a manager to be given latitude to try out new ideas in the new setting. Through formal and informal communication, these managers will often reflect new knowledge gained through such

[3]The biggest advantage of studying hotels using Economic Census data over using the dataset from Smith Travel—the most commonly used dataset for analyzing the hotel industry—is that the Economic Census provides data on revenues and payroll so that one can estimate productivity effects. The Economic Census is also more complete, as it contains every hotel in the United States, while Smith Travel only has data on larger chain hotels. However, the Smith Travel dataset has more complete information on brands and franchising than the Economic Census.

[4]Labor factor productivity is measured as the residual from a pooled OLS regression of log revenue on log payroll with year fixed effects.

experiments back to their former colleagues in the pre-existing hotel. Thus, the experience of industry experts supports the idea that knowledge transfers may be an important source of internal agglomeration effects.[5]

In our empirical specifications, we define a market to be a metropolitan statistical area (MSA). This definition of a market represents an imperfect compromise between defining markets as areas large enough to be truly distinct, while small enough to meaningfully capture a local area for the purposes of defining entry and exit. Other studies of internal agglomeration in the hotel industry have typically used smaller market definitions, particularly in studies focusing on resource sharing, such as referrals (Ingram & Roberts, 2000) or physical asset sharing. Our MSA-based market definition allows us to exploit geographic variation within-MSA to help distinguish between within-county resource sharing effects and broader market effects that obtain within-MSA but between counties.

For our main results, we rely on difference-in-differences estimators—within- and between-establishment changes in productivity from changes in the intensity of internal agglomeration, as measured by the count of establishments owned by the same firm in the same market. Establishments that appear in only 1 year, those outside of metropolitan statistical areas, and establishments that are local monopolies are dropped from the sample, though the results are not sensitive to including these observations. The resulting dataset used for the main set of hotel results contains approximately 59,000 establishments and 130,000 establishment-years from about 24,000 firms.[6]

3.2 | Internal agglomeration and performance

In our baseline tests, we estimate the impact of changes in the intensity of internal agglomeration on establishment-level changes in productivity. Specifically, for establishment i infirm j, market (MSA) g and year t we regress labor factor productivity, *LFP*, on the log of the number of establishments from the same firm in the same market, *ESTABS*, with establishment, λ, and year, T, fixed effects and vector of controls $\boldsymbol{X}$ indexed by c:

$$LFP_{it} = \alpha + \beta_1 ESTABS_{jgt} + \lambda_i + \boldsymbol{T_t} + \boldsymbol{X_{cit}B_c} + \varepsilon_{it} \quad (1)$$

In expression (1), *LFP* is measured as the establishment-level residual from a cross-sectional (year-by-year) regression of log sales on log payroll, and $ESTABS_{jgt}$ is the count (or log count) of the number of establishments owned by a firm in a particular market at a point time ("firm-MSA establishments" in the results and tables below).[7] The count of the number of

[5]The information in this section came from 20 interviews with hotel and restaurant industry experts. We are particularly grateful to Phil Borkowski (Marriott), Krissy Gathright (Apple REIT), Stephen Pettibone (Starwood), and Adam Weissenberg (Deloitte), for sharing their wealth of experiences with us.

[6]The U.S. Census Bureau requires us to report rounded numbers of observations in our sample.

[7]There are several, roughly equivalent, methods of measuring productivity changes including one-step, two-step, and first-differencing. In our two-step estimation, establishment productivity is first estimated as the residual from a period-by-period cross-sectional regression of log revenue on log payroll, across all establishments, and then used as the dependent variable in a differences-in-differences regression with establishment and time fixed effects. In one-step productivity estimation log revenue is regressed on log payroll along with the key explanatory variables and controls. In a first differences specification, one takes within-establishment differences in all the covariates period by period and then regresses changes in productivity on changes in the intensity of internal agglomeration. Our two-step productivity results are nearly identical to both alternative specifications.

establishments (*ESTABS*) is a good measure of the intensity of internal agglomeration, because it captures the physical scope of a firm's presence within a market, as measured by discrete operating units. Thus, *ESTABS*, which measures the intensity of internal agglomeration, can be interpreted as a productivity shift parameter.[8]

The vector $\boldsymbol{X}$ contains variables that proxy for other potential shifters of an establishment's productivity frontier, including: log establishment age ("log age"), measured as the log of the difference between the observation year and the year in which the establishment was first captured in the Economic Census; the log of the count of competitors' establishments in the same market ("log MSA establishments"); a Herfindhal-Hirschman index of the concentration of sales amongst the 10 largest establishments within a market ("HHI"); the average productivity of other establishments in the same market ("MSA productivity"); an ownership change dummy that is equal to one when the focal establishment reports being owned by a different entity at time t compared to time $t-1$ and zero otherwise; the log of one plus the number of establishments owned by the same firm in other markets ("log firm establishments ex-MSA"); a lagged firm productivity spline (quartiles), measured as the average revenue to payroll ratio of all the firm's other establishments in the prior period ("lagged firm sales multiplier quartiles"); and hundreds of market-year fixed effects. All dollar-denominated values are reported in 2019 dollars using CPI deflators from the Bureau of Labor Statistics. ϵ_{it} is a mean-zero, normally distributed error term. Standard errors are clustered at the establishment-level, though the results are robust to clustering at the firm or market level as well. Notably amongst the controls, log count of competitors' establishments in the same market, both controls for interfirm agglomeration effects as well as competition effects. Market-year fixed effects are also important, as they control for all time-varying market-level variation that might be correlated with internal agglomeration decisions and productivity (e.g., supply and demand shocks), ruling out a large set of confounds.

Expression (1) decomposes productivity into components, including an establishment-specific component, and a time-varying within-establishment component. The data are rich, and the controls extensive, but the specification does not include two variables of interest. First, brand information is not reported to the Census Bureau so our estimate of the effect of internal agglomeration should be interpreted as firm-level, not brand-level, agglomeration effects. For example, 48% of Host Hotels' revenue comes from the 54 Marriott hotels it owns, while 52% comes from the 60 other branded hotels it operates, including 9 Hyatts, 13 Westins, and 8 Sheratons.[9] Since the establishment-level effects we estimate are determined by the owner (e.g., Host Hotels) and not by the brand (e.g., Sheraton), and hotels and restaurants are often franchised, the absence of brand information means that there will be markets where two establishments of the same brand are co-located but are owned by two different franchisees.

To the extent that brand agglomeration effects are meaningful (Ingram & Baum, 1997), our firm-level results will tend to underestimate the "true" agglomeration effect in two ways. First, by confining our estimates to within-firm agglomeration we miss positive internal agglomeration effects that operate at the brand, but not the firm-level. Second, by comparing internal agglomeration "treatments" against control group establishments that benefit from brand

[8]Our measure of internal agglomeration—*ESTABS*—is the count of other establishments owned by the parent firm in the MSA. Other potential measures include counts by other geographic units or by distance, or by number of employees, or by counts within prespecified travel time. We choose an MSA-level measure because interviews indicated MSA to be the relevant geographic unit to capture knowledge sharing between jointly owned hotels. MSA is also a relatively standard measure of a market, allowing for easy comparison across industries and time.

[9]Host Hotels 2014 10-K. Accessible at http://quicktake.morningstar.com/stocknet/secdocuments.aspx?symbol=hst.

agglomeration effects, we would overestimate the baseline against which internal agglomeration effects are measured, and therefore underestimate internal agglomeration effects (i.e., because we have a differences-in-differences specification). While it would be interesting to capture both firm and brand-level agglomeration effects, the former appears to be especially understudied in the literature. And yet, one might expect within-firm agglomeration effects to be quite important, even when a firm operates multiple brands. To address the issue of missing brand data directly empirically, we use information in the 2007 Census data, where franchising data is available, to show in robustness tests (Table 9, column 3) that our results hold even when controlling for franchised establishments (i.e., where omitted brand effects are particularly important).

Second, we do not observe investment flows. Cross-sectional variation in fixed capital stock is controlled for directly with establishment-level fixed effects. However, if capital expenditures, or investments in brand building (e.g., advertising), are systematically larger (smaller) at preexisting establishments when the firm increases (decreases) the number of establishments it owns in the same market, our productivity estimates would be biased upward (downward) due to omitted variable bias. Fortunately, capital flows and brand building expenditures should be highly correlated with labor flows and several variables in $\boldsymbol{X_c}$, including: establishment age, secular trends, average MSA productivity, and local market by year fixed effects. Furthermore, while initial capital expenditures are substantial, the data suggest that ongoing investment flows are relatively small compared to labor flows for hotels and restaurants. Thus, any residual bias from including proxies for capital flows, instead of capital flows themselves, is likely to be of small magnitude economically.[10]

3.3 | Empirical strategy

In the ideal experiment, we would randomly change the intensity of internal agglomeration and measure the resulting productivity effects directly. Greenstone et al. (2010) exploit quasi-experimental variation in external agglomeration, comparing "winning" counties that lobbied for so-called "Million Dollar Plants" to one or two "losing" runner-up counties that survived a long selection process but narrowly lost to the winner. However, such an approach is not feasible for our study of internal agglomeration, as firms do not systematically disclose the geographic areas in which they are considering co-locating.

Given that our data is not generated by an experimental process, expression (1) can estimate the conditional correlation between internal agglomeration and performance but, one can only interpret the correlation as a causal effect of internal agglomeration on productivity if there are no omitted variables in the specification that are correlated with the dependent variable and the error term. In other words, the identifying assumption in expression (1) is that there are no factors missing in our difference-in-differences specification that would both increase (decrease)

[10]For example, according to Lodging Magazine's analysis of the trade report, the "Uniform System of Accounts for the Lodging Industry," labor costs average around 1/3 of revenue http://lodgingmagazine.com/examining-hotel-labor-costs/. (Last accessed on May 22, 2015.) By comparison, Host and Hersha, two of the largest publically listed hotel-focused equity real estate investment trusts (equity REITs are property owner-operators) spent approximately 9% of their revenue on capital expenditures. And, a Cornell Hospitality School survey of 2,815 U.S. hotels reported non-payroll marketing expenditures to be only about 4% of sales (O'Neill, Hanson, & Mattila, 2008). Host's 10-K is available at http://quicktake.morningstar.com/stocknet/secdocuments.aspx?symbol=hst. Hersah's 10-K is available at: http://www.snl.com/Cache/c27840380.html.

the productivity of the focal establishment and increase (decrease) an establishment's intensity of internal agglomeration. While it seems plausible that the coefficient on the intensity of internal agglomeration β_1 is well identified in expression (1)—it controls for time-invariant establishment-specific factors, market level trends, and proxies for capital flows—it is possible that non-random selection processes, omitted variable bias or heterogeneous treatment effects might bias our estimates. For example, managers may systematically internally agglomerate when a pre-existing establishment is just about to experience a positive productivity boost, or where internal agglomeration benefits are expected to be largest. Another concern may be that an increase in the intensity of internal agglomeration is correlated with new investment in markets that benefit pre-existing firms. To deal with the potential for endogeneity bias due to selection effects, we use matching on observables to deal with selection on observable *ex ante* differences between establishments, and an instrumental variables technique to address selection on unobservable (to the econometrician) characteristics of firms and establishments.

Matching is used to identify control group establishments that "look like" treated establishments prior to the change in the intensity of internal agglomeration to control for the possibility that selection on observables is biasing the results. Because the processes firms undertake to open a new establishment are likely different from those undertaken to close an existing establishment, we create two different matched samples in the first stage—one for each type of treatment (positive or negative)—and then combine the two matched samples in a second stage analysis, using expression (1). Specifically, using Coarsened Exact Matching (CEM), developed by Iacus, King, and Porro (2011), we specify two sets of many-to-one matches, predicting separately: (a) increasing intensity of internal agglomeration, $TREAT+$, ("positive treatment"), and (b) decreasing the intensity of internal agglomeration, $TREAT-$, ("negative treatment"), using:

$$TREAT+_{it} = f(l_{it-1}, \boldsymbol{T_t}, \boldsymbol{X_{cit-1}}), \tag{2}$$

$$TREAT-_{it} = f(l_{it-1}, T_t, \boldsymbol{X_{cit-1}}), \tag{3}$$

where $\boldsymbol{X_{cit\text{-}1}}$ includes all of the controls from expression (1) lagged one period (i.e., 5 years, given the nature of the Census data) and coarsened into deciles when continuous, plus two measures of lagged productivity: the first a measure where the components of lagged productivity—lagged sales and lagged payroll—enter separately, and the second, perhaps more intuitive productivity control, lagged establishment-level sales divided by payroll.[11]

Once the two matched samples are identified (i.e., from (2) and (3)), they are combined and analyzed using expression (1). If any remaining unobservable establishment-specific time-varying information does not affect both the internal agglomeration selection process and changes in labor factor productivity at the focal establishments, then the matched sample will be equivalent to random assignment of internal agglomeration to legacy establishments. That is, if the matched sample represents a valid counterfactual to the treatment group, one can interpret the coefficient on internal agglomeration intensity as the average treatment effect of internal agglomeration.

[11]Firm sales multipliers were already lagged in expression (1), they were not lagged an additional period in the matching algorithm. Although Census disclosure rules do not permit us to tabulate the first stage results, the fact that CEM matches treatment to control observations exactly, suggests that one need not evaluate the first stage for balance checking, as in propensity score matching.

To address concerns that unobservable time-varying establishment-specific factors might confound the matched sample results, we follow the dynamic panel data estimation literature (Holtz-Eakin, Newey, & Rosen, 1988), by specifying expression (1) in first differences and using lagged levels "internal" right hand side variables as instruments for the key endogenous variable (i.e., *COLO*). Dynamic panel data estimators are considered to be particularly appropriate to contexts, such as ours, with a large number of observations per time period, and relatively few time periods.

Although dynamic panel data methods are widely used (Arellano & Bover, 1995; Blundell & Bond, 1998), and broadly accepted in strategic management (Bettis, Gambardella, Helfat, & Mitchell, 2014), the "mechanical" nature of the estimator—the exclusion restriction is satisfied by functional form—renders them "a less good solution" (p. 951) to many. To address such reasonable concerns about dynamic panel data estimation techniques, we chose an instrument that could plausibly satisfy the exclusion restriction assumption conceptually as well as mechanically. That is, we have a priori reasons to believe *lagged log firm establishments ex-MSA* should have no direct influence on changes in a focal establishment's productivity, conditional on the broad set of controls we employ (e.g., including establishment fixed effects and changes in firm productivity and *log firm establishments ex-MSA*). Moreover, using a single instrument obviates the well-known "many weak instruments" problem of dynamic panel data estimators (Bun & Windmeijer, 2010).

While the fact that one cannot test the exclusion restriction directly is also unsatisfying, the identifying assumption we follow does seem credible in our context, particularly because the baseline results show no conditional correlation between contemporaneous *log firm establishment ex-MSA* and *productivity*. (To ensure this assumption is not violated by establishment-level heterogeneity, we include establishment fixed effects in both stages of the 2SLS estimator as well.) And yet, the instrument *lagged log firm establishments ex-MSA* should be strong (i.e., correlated with changes in *COLO*) if firm size outside of a focal market is broadly predictive of size within a focal market—an assumption we test directly. While we remain agnostic as to whether "mechanical" identification is better or worse than conceptual identification—both are clearly "second best" compared to experimental variation—it is reassuring when standard endogeneity corrections yield results that are consistent with differences-in-differences and matched sample results.

4 | RESULTS

4.1 | Summary statistics

The description of variables and summary statistics are presented in Tables 1 and 2, respectively. Census rules require that we do not report precise distributional statistics such as maximums, minimums or medians, however, we can report the mean and standard deviation of the distributions, as well as the top and the bottom decile. The summary statistics indicate that firms and establishments in our sample are relatively small with mean sales of $4 million. While the modal firm operates a single hotel, multiunit firms average 32 hotels (~3 within a focal MSA and ~ 29 outside the MSA). Though the results are not sensitive to restricting the sample to larger firms and establishments, it is useful to note that the typical firm in our sample is not a global hospitality firm, like the Marriott Corporation, but rather a small to mid-sized firm.

TABLE 1 Key variable descriptions

Variable name	Description
Productivity	Labor factor productivity for establishment *i* in year *t*, estimated as the residual from the period by period pooled cross-sectional regression: $log\ rev_i = a + b*log\ payroll_i + e_i$. ("*LFP*")
Sales	Total sales for establishment *i* in year *t*
Payroll	Total payroll for establishment *i* in year *t*
Firm-MSA establishments	For establishment *i*, count of establishments from same parent firm *j* in the same MSA *g* in year *t*. ("*ESTABS*")
Age	The number of years establishment *i* has been in the dataset in year *t*
Firm establishments ex-MSA	For establishment *i*, the count of establishments with same parent firm *j* in other MSAs (not *g*) in year *t*
MSA establishments	For establishment *i*, the count of all other establishments in the same MSA *g* in year *t*
MSA productivity	For establishment *i*, average productivity of all other establishments in same MSA *g* in year *t*
Firm sales multiplier	$=sales_{it}/payroll_{it}$
Positive treatment	A dummy variable equal to one when firm *j* increases the number of establishments in establishment *i*'s MSA *g*in year *t*, and zero otherwise. ("*TREAT+*")
Negative treatment	A dummy variable equal to one when firm *j* decreases the number of establishments in establishment *i*'s MSA *g* in year *t*, and zero otherwise. ("*TREAT−*")

Table 2 shows that the intensity of internal agglomeration was relatively low with an average of 1.54 establishments owned by the same firm in an MSA (*Firm-MSA establishments*). Over 50% of the positive treatments, or increases in internal agglomeration, we observe are changes from one to two or two to three hotels in the same market owned by the same firm. Similarly, most of the negative treatments, or reductions in internal agglomeration, were changes from two to one or three to two hotels. Positive and negative treatments each occurred in about 2% of the observations. The distribution of our key explanatory variable (*Firm-MSA establishments*) is skewed, with tens of thousands of observations having no collocated establishments from the same firm in their market.

Table 3 shows the pairwise correlation coefficients between the variables in Table 2. Only five correlations are greater than 0.4. Since *productivity* is estimated from a regression of *sales* on *payroll*, and the *sales multiplier* is *sales* divided by *payroll* it is not surprising that these two variables are correlated at 0.61. Given that many productivity studies have shown a high correlation between revenue and labor, and that several industry experts noted the close relationship between changes in *sales* and changes in *payroll* in the hotel industry, it is also not surprising that these variables are correlated at 0.98. The high correlation also suggests that our one-factor productivity measure of performance is a strong proxy for multifactor productivity, since changes in payroll are probably highly correlated with changes in capital expenditures. *Firm-MSA establishments* is correlated with three other variables of note. *Positive treatment* and *Firm-MSA sales*$_{-i}$ are mechanically related to *Firm-MSA establishments* (correlations of 0.48 and 0.44, respectively), since a firm must have at least two establishments in an MSA to be positively

TABLE 2 Summary statistics ($n \sim 130{,}000$)

Variable	Mean	Stdv	Mean of bottom decile	Mean of top decile
[1] Productivity	0.10	0.44	−0.72	0.89
[2] Sales multiplier	3.57	4.74	2.13	17.75
[3] Sales ($000)	3,999	19,750	174	31,200
[4] Payroll ($000)	1,119	6,033	19	9,516
[5] Firm-MSA establishments	1.54	2.16	1	7
[6] Positive treatment	0.02	0.15	0	1
[7] Negative treatment	0.02	0.12	0	1
[8] Firm-MSA sales$_{-i}$	3,245	25,930	0	29,760
[9] Age (years)	15.3	5.4	10	27.5
[10] Ownership change	0.05	0.21	0	1
[11] MSA establishments	336	272	38	977
[12] Concentration index	0.04	0.64	0.01	0.19
[13] MSA productivity$_{-i}$	0.00	0.10	−0.18	0.21
[14] Firm establishments ex-MSA| >0	29.3	116.7	1.5	n/a
[15] Firm sales multiplier	4.45	0.88	2.52	7.94
[16] Year	1998	6.4	1987	2007

Note: The Census Bureau's current disclosure rules do not allow the reporting of variable maximum and minimum values, so we report the means of the bottom decile (the 1st to 10th percentile) and top decile (the mean of the 90th–99th percentile). We report descriptive statistics for *Firm establishments ex-MSA* for those establishments in firms with at least two establishments. Most establishments are in firms with one establishment. Census disclosure rules preclude us from reporting the mean of the top decile of the distribution of this (conditional) variable, since doing so might reveal information on an individual firm.

treated and to have sales in the MSA that are not from the focal establishment. The correlation (0.53) between *Firm-MSA establishments* and *Firm establishments ex-MSA*, suggests that firms with hotels in multiple markets are more likely to internally agglomerate than firms operating in a single market.

4.2 | Main results

We investigate the effect of within-establishment changes in the intensity of internal agglomeration on establishment performance in Table 4. In Column 1, we report a pooled cross-sectional regression with time fixed effects. The relationship between the *Firm-MSA establishments* and *productivity* is positive and precisely estimated, meaning that more intensively agglomerated establishments are more productive than establishments that are less intensively agglomerated.

In Columns 2 and 3, we run difference-in-differences estimators with establishment fixed effects. Establishment fixed effects absorb all time-invariant establishment-specific sources of variation in productivity, including any differences due to cross-sectional variation in capital stock, location, and brand. In these specifications, the coefficient on *Firm-MSA establishments* is 0.004 and precisely estimated, and the R^2 increases from 0.03 to 0.72. The interpretation is

TABLE 3 Pairwise correlations ($n \sim 130{,}000$)

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]
[1] Productivity															
[2] Sales multiplier	0.61														
[3] Sales	0.22	−0.08													
[4] Payroll	0.18	−0.09	0.98												
[5] Firm-MSA estabs	0.12	−0.09	0.03	0.02											
[6] Positive treatment	0.08	−0.02	0.04	0.03	0.48										
[7] Negative treatment	0.03	−0.04	0.03	0.03	0.14	−0.02									
[8] Firm-MSA $sales_{-i}$	0.12	−0.04	0.20	0.18	0.44	0.22	0.11								
[9] Age	0.00	−0.09	0.04	0.04	0.06	0.08	0.09	0.04							
[10] Ownership change	0.10	−0.04	0.06	0.06	0.17	0.09	−0.02	0.08	0.14						
[11] MSA establishments	0.07	−0.02	0.07	0.07	0.15	0.07	0.05	0.10	0.10	0.03					
[12] Concentration index	−0.02	−0.00	−0.02	−0.02	−0.06	−0.03	−0.01	−0.05	0.00	0.01	−0.34				
[13] MSA $productivity_{-i}$	0.04	0.03	−0.01	−0.02	0.04	0.03	0.01	0.01	−0.02	0.02	0.07	0.14			
[14] Firm estabs ex-MSA	0.12	−0.03	−0.00	−0.01	0.53	0.33	0.13	0.11	0.11	0.17	0.01	0.01	0.03		
[15] Firm sales multiplier	0.14	0.20	−0.05	−0.06	0.16	0.10	0.01	−0.03	−0.00	0.03	−0.01	0.00	0.02	0.26	
[16] Year	0.12	−0.05	0.01	0.01	0.07	0.06	0.07	0.03	0.56	0.15	0.18	0.01	−0.01	0.07	0.01

TABLE 4 Internal agglomeration and performance: baseline results

	(1)	(2)	(3)	(4)	(5)	(6)
Specification	OLS	D-in-D	D-in-D	D-in-D	D-in-D	Matched
Dep. variable=	productivity	productivity	productivity	productivity	sales mult.	productivity
Firm-MSA establishments	0.023	0.004	0.005	0.005	0.033	0.005
	(0.001)	(0.001)	(0.001)	(0.001)	(0.007)	(0.002)
Log age			0.0108	0.116	−0.757	−0.033
			(0.012)	(0.009)	(0.230)	(0.024)
Log MSA establishments			−0.007	−0.025	−0.152	−0.029
			(0.007)	(0.009)	(0.100)	(0.015)
Concentration index			−0.163	−0.143	−0.194	−0.127
			(0.033)	(0.030)	(0.306)	(0.054)
MSA productivity			0.139	0.136	0.507	0.130
			(0.018)	(0.016)	(0.186)	(0.027)
Log firm establishments ex-MSA			−0.000	0.000	−0.004	0.003
			(0.003)	(0.002)	(0.012)	(0.03)
Year fixed effects	Y	Y	Y	Y	Y	Y
Firm sales multiplier quartiles	N	N	Y	Y	Y	Y
Ownership change dummy	N	N	Y	Y	Y	Y
Establishment fixed effects	N	Y	Y	Y	Y	Y
MSA × year fixed effects	N	N	N	Y	Y	Y
n	~130,000	~130,000	~130,000	~130,000	~130,000	~57,000
Adjusted R^2	0.03	0.72	0.72	0.77	0.72	0.55

Note: Standard errors clustered at the establishment level.

that increasing the intensity of internal agglomeration by one unit is associated with a 0.4% increase in productivity, even after controlling for establishment-specific differences. Adding additional controls in Column 3—including controls for MSA size, concentration and productivity; firm size and performance; and establishment age and ownership changes—does not meaningfully change the magnitude or precision of the key explanatory variable. Column 4—our "main result"—reports the results including a full set of MSA × year fixed effects, which control for a wide range of time-varying market-specific factors, such as changes in market attractiveness due to changes in demand and labor supply. The inclusion of MSA × year fixed effects improves the adjusted R^2 to 0.77, but does not change the magnitude or precision of the coefficient on *Firm-MSA establishments*.

Columns 5 replicates Columns 4, but with *sales multiplier* as the dependent variable instead of *productivity*. The point estimate on *Firm-MSA establishments* is 0.033 (*t*-statistic = 4.7). The

interpretation is that an increase in the intensity of internal agglomeration by one unit leads to a 0.9% (=0.033/3.57) in an establishment's *sales multiplier*.

Column 6 replicates Column 4 after using our matching approach. This column tabulates the "second stage" of the CEM matched sample regressions after eliminating control group observations that do not match to treatment group observations (and vice versa). The point estimate on *Firm-MSA establishments* is 0.005 (t-statistic = 2.5), which is similar in magnitude as in Column 4. The result is important because it suggests that selection on observables does not have a large impact on the magnitude of internal agglomeration effects in the hotel sample. A typical "treated" hotel is different from many untreated hotels *ex ante*, but such heterogeneity does not bias the results upward.

Coefficients on several other variables are interesting as well. The coefficient on *MSA-establishments* is negative across all specifications. The inclusion of this variable controls for the net effect of competition and external agglomeration. The coefficient estimate suggests competition effects swamp external agglomeration effects in the hotel industry. The coefficient on *MSA productivity* is positive across all models. The inclusion of this variable controls for the effect of overall productivity changes in the focal establishment's market, demonstrating that market-level productivity is positively correlated with a focal establishment's productivity.

Taken together, the results from Table 4 suggest a positive relationship between internal agglomeration and performance, even after controlling for time-invariant establishment specific characteristics, and a host of time-varying establishment, firm and market characteristics, as well as for selection on observables.

In Table 5, we provide a number of results using alternative explanatory variables. These serve as robustness tests but also help us to determine which mechanisms are more or less plausible. In Column 1, we use *log Firm-MSA establishments* as the dependent variable. The coefficient on this term of 0.024 (t-statistic = 3.9) provides a useful economic interpretation as it captures the productivity effect of internal agglomeration as the intensity of internal agglomeration doubles, as measured by the number of internally agglomerated units in an MSA. The interpretation is that doubling the intensity of internal agglomeration is associated with a productivity increase of about 2% in pre-existing establishments. Comparing the magnitude of the effect as internal agglomeration intensity doubles to the magnitude of the effect when internal agglomeration intensity increases by one establishment (regardless of how many establishments were internally agglomerated previously), as in column 4 of Table 4, is instructive. The fact that the effect size is almost 50 times larger when measured as a doubling of internal agglomeration intensity, compared to increasing internal agglomeration intensity by one unit, tells us that there are diminishing marginal returns to co-locating units.

In Column 2, we use *log Firm-MSA sales*$_{-i}$ (the log sum of all sales by the parent firm in the MSA, excluding the sales of the focal establishment) as an alternative measure of internal agglomeration intensity. The coefficient on this term is 0.005 (t-statistic = 4.9), suggesting that doubling the intensity of internal agglomeration, as measured by sales, leads to a 0.5% increase in productivity. The result provides evidence that the internal agglomeration and productivity result is robust to alternative measures of internal agglomeration intensity.

In Column 3 we split *Firm-MSA establishments* into two mutually exclusive explanatory variables, based on the fact that counties are always subsets of MSAs: *Firm-MSA establishments-Firm-county establishments* and *Firm-county establishments*. These variables are counts of the number of sister establishments in the same MSA but not the same county, and counts of the number of sister establishments in the same county, respectively. Since county is a smaller geographic area than MSA, one can think of these two variables as capturing the effect from

TABLE 5 Alternative explanatory variables

Dep. variable = *productivity*	(1)	(2)	(3)	(4)	(5)
Log firm-MSA estabs	0.024				
	(0.006)				
Log firm-MSA $sales_{-i}$		0.005			
		(0.001)			
Firm-MSA estabs—firm-county estabs			0.007		
			(0.001)		
Firm-county estabs			0.002		
			(0.002)		
Log(firm-MSA estabs—firm-county estabs)				0.029	
				(0.008)	
Log firm-county estabs				0.014	
				(0.007)	
1–3 firm-MSA estabs					0.029
					(0.008)
4+ firm-MSA estabs					0.038
					(0.008)
Controls	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y
Firm sales multiplier quartiles	Y	Y	Y	Y	Y
Own. change dummy	Y	Y	Y	Y	Y
Establishment fixed effects	Y	Y	Y	Y	Y
MSA × year fixed effects	Y	Y	Y	Y	Y
n	~130,000	~130,000	~130,000	~130,000	~130,000
Adjusted R^2	0.77	0.55	0.56	0.56	0.77

Note: Standard errors are clustered at the establishment level.

"proximate but not adjacent" internal agglomeration and "adjacent" internal agglomeration, respectively.

The coefficient on the first variable is 0.007 (*t*-statistic 6.8), the coefficient on the second variable is 0.002 (*t*-statistic 1.0), and the *t*-statistic on the difference between them is 2.2. The result suggests that while internal agglomeration affects productivity, the internally agglomerated establishments do not need to be located adjacent to each other in a "cheek-by-jowl" fashion to generate positive productivity effects. One further implication might be that internal agglomeration is even more important when establishments are "near but not too near" one another. However, we caution that the result is not robust to specification. For example, we repeat the analysis in Column 4 but use log counts of these two variables instead of levels. The coefficients are still positive, but the *t*-statistic on the difference is only 1.4, even though the coefficient on *log*(*Firm-MSA establishments-Firm-county establishments*) is about twice the magnitude of the coefficient on *log Firm-county establishments*. Nevertheless, the results in Columns 3 and 4 help to partially rule out the idea that investment in shared physical assets are the primary

TABLE 6 2SLS

	(1)	(2)
	1st stage	**2nd stage**
Dep. variable=	ΔFirm-MSA estabs	Δproductivity
Lag log firm establishments ex-MSA	0.001	
	(0.000)	
ΔFirm-MSA establishments		0.068
		(0.011)
Δlog age	0.081	0.116
	(0.024)	(0.010)
Δlog MSA establishments	0.147	−0.018
	(0.032)	(0.007)
ΔConcentration index	0.019	−0.154
	(0.056)	(0.031)
ΔMSA productivity	−0.016	0.155
	(0.031)	(0.017)
Δlog firm establishments ex-MSA	0.592	−0.042
	(0.033)	(0.008)
Year fixed effects	Y	Y
ΔFirm sales multiplier quartiles	Y	Y
Ownership change dummy	Y	Y
Establishment fixed effects	Y	Y
MSA × year fixed effects	Y	Y
n	~130,000	~130,000
Adjusted R^2	0.17	n/a
F-test on the instrument	161	

mechanism driving the internal agglomeration result in our setting. If internal agglomeration resulted in investment in shared physical resources (e.g., a single catering kitchen that is used by two adjacent hotels), then we would expect a stronger result on "adjacent" internal agglomeration compared to "proximate" internal agglomeration, but there is no evidence of such an effect.

In Column 5, we introduce dummies for whether the focal establishment operated in an MSA with 1–3 additional sister establishments (*1–3 Firm-MSA establishments*) or 4 or more additional sister establishments (*4+ Firm-MSA establishments*), categorical cut-offs that roughly correspond to the median number of internally agglomerated establishments, conditional on being internally agglomerated, in our sample. The coefficients on each variable are positive and similar to each other in magnitude (we cannot reject that the coefficients are equal). The results suggest that being internally agglomerated with a "small" number of establishments or a "large" number of establishments, results in, more-or-less, the same average benefit to the focal establishment.

To challenge the causal interpretation of our results further, we conduct a two-stage least squares (2SLS) regression. The approach helps address concerns that unobservable time-varying establishment-specific factors might confound our main or matched sample results. These results are presented in Table 6. Column 1 presents the "first stage" results. The coefficient on our instrument, *lagged log firm establishments ex-MSA*, is positive with a *t*-statistic of 14.7 and an *F*-statistic of 161, well above the value of 10 suggested by Staiger and Stock (1997). Thus, the instrument is strong. Column 2 presents the "second stage" results. The coefficient on the (instrumented) *Firm-MSA establishments* variable is positive (*t*-statistic = 6.1), which is consistent with our main results. While the economic interpretation of the 2SLS estimator is somewhat different than the difference-in-differences estimator—the former captures the change in the increase in productivity due to changes in increased internal agglomeration intensity, while the latter captures increases in productivity due to increased internal agglomeration intensity—the result, nevertheless, provides some additional support for a causal interpretation of the internal agglomeration and productivity results in Tables 4 and 5.

4.3 | Mechanisms

Beyond showing that internal agglomeration causes productivity to increase, we want to understand the mechanisms behind internal agglomeration effects. As a first step, we disentangle the contemporaneous and persistent effects of internal agglomeration (with respect to state—internally agglomerated or not) by discretizing changes in internal agglomeration intensity into positive and negative categorical variables, as in specification (4) above.

In Table 7, we show difference-in-differences and matched sample productivity regressions, where the key explanatory variables are *positive treatment* and *negative treatment*. *Positive treatment* captures all the net benefits of internal agglomeration, including both contemporaneous and persistent effects. *Negative treatment*, however, only picks up a reduction in the net contemporaneous effects of internal agglomeration, as any persistent effects (with respect to internal agglomeration state), by definition, would have already been absorbed into the production functions of the remaining establishments in the local market. Thus, by comparing the coefficient estimates on positive and negative internal agglomeration we can separately recover contemporaneous and persistent effects of internal agglomeration on productivity.[12]

In Column 1, the difference-in-differences estimator results in *positive treatment* being positively correlated with productivity. In Column 2, the difference-in-differences estimator results in *negative treatment* being negatively correlated with productivity. The interpretation of the coefficient estimate of 0.033 on *positive treatment* is that increasing the intensity of internal agglomeration is associated with increasing productivity by 3.3%; the coefficient estimate of −0.023 on *negative treatment* means that decreasing the intensity of internal agglomeration is associated with a 2.3% decline in productivity.

In Column 3, *positive treatment* and *negative treatment* enter together. The coefficient estimates are 0.030 and −0.016, respectively, and the *t* test on the difference is 1.75. These results suggest that the contemporaneous benefit of internal agglomeration is approximately 1.6% and the net persistent effect of internal agglomeration is approximately 1.4% (0.030–0.016).

[12]Since census data is collected every 5 years, persistent benefits are measured over a relatively long period of time. For example, if the "treatments" are uniformly spaced over time, the average time between the treatment and outcome would be 2.5 years.

TABLE 7 Positive and negative treatment effects

Dep. variable = *productivity*	(1)	(2)	(3)	(4)
Specification	D-in-D	D-in-D	D-in-D	Matched
Positive treatment	0.033		0.030	0.032
	(0.006)		(0.006)	(0.009)
Negative treatment		−0.023	−0.016	−0.017
		(0.007)	(0.007)	(0.010)
Log age	0.115	0.116	0.115	0.115
	(0.009)	(0.009)	(0.09)	(0.09)
Log MSA establishments	−0.025	−0.025	−0.025	−0.025
	(0.009)	(0.009)	(0.009)	(0.009)
Concentration index	−0.142	−0.142	−0.142	−0.142
	(0.030)	(0.030)	(0.030)	(0.030)
MSA productivity	0.136	0.136	0.136	0.136
	(0.016)	(0.016)	(0.016)	(0.016)
Log (firm establishments ex-MSA + 1)	0.002	0.004	0.003	0.003
	(0.002)	(0.002)	(0.002)	(0.002)
Year fixed effects	Y	Y	Y	Y
Firm sales multiplier quartiles	Y	Y	Y	Y
Ownership change dummy	Y	Y	Y	Y
Establishment fixed effects	Y	Y	Y	Y
MSA × year fixed effects	Y	Y	Y	Y
t test on difference			1.75	1.21
Pos Treat − |Neg Treat|				
n	~130,000	~130,000	~130,000	~57,000
Adjusted R^2	0.77	0.77	0.77	0.55

Note: Standard errors clustered at the establishment level.

Column 4 tabulates the second stage of the CEM matched sample regressions. The point estimate on *positive treatment* is 0.032 and the point estimate on *negative treatment* is −0.017. However, the *t* test on the difference between the two is 1.21. While these results are consistent with those in Column 3, the loss in precision may be a result of the decrease in the number of observations. We interpret the results from Table 7 cautiously as providing suggestive evidence that persistent benefits of internal agglomeration (with respect to internal agglomeration state) are at least as important as the contemporaneous effects of internal agglomeration.

To further explore the internal agglomeration effect, we perform a number of extensions and robustness checks. First, we proxy for the importance of improved monitoring as a mechanism, by studying whether internal agglomeration effects are marginally weaker when all of a firm's establishments are in the same market. Assuming that a firm's headquarters can be traced to a market when all of its operations are in that market—an assumption that seems natural, particularly for small and mid-sized hotel firms—one can devise a simple test of the marginal effect of improved monitoring on internal agglomeration, by comparing internal

TABLE 8 Extensions and robustness checks

Dep. variable = *productivity*	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Mechanism or robustness test	Monitoring	Reflection	Placebo	Placebo	Franchise
Firm-MSA establishments	0.004	0.003	−0.003		0.007
	(0.001)	(0.001)	(0.001)		(0.001)
All hotels in one MSA	0.034				
	(0.014)				
Firm-MSA establishments × All hotels in one MSA	0.005				
	(0.003)				
Firm-MSA establishments × MSA productivity		0.039			
		(0.005)			
Positive treatment				0.006	
				(0.007)	
Negative treatment				−0.004	
				(0.009)	
Firm-MSA establishments × Franchisee					−0.009
					(0.003)
Firm-MSA establishments × Franchisor					−0.006
					(0.003)
Controls	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y
Firm sales multiplier quartiles	Y	Y	Y	Y	Y
Own. change dummy	Y	Y	Y	Y	Y
Establishment fixed effects	Y	Y	Y	Y	Y
MSA × year fixed effects	Y	Y	Y	Y	Y
n	~130,000	~130,000	~84,000	~84,000	~130,000
Adjusted R^2	0.77	0.56	0.77	0.77	0.77

Note: In the placebo tests, all the variables, except firm-MSA establishments (Column 3) and the positive and negative treatment dummies (Column 4), are lagged one period.

agglomeration effects in single market firms to multi-market firms. Recall that the prior literature suggests that monitoring costs tend to increase with distance between an establishment and its headquarters, but that those costs can be ameliorated by internal agglomeration. If the prior literature is correct, then an internally agglomerated establishment in a firm that operates in multiple markets should enjoy a larger productivity increase compared to an internally agglomerated establishment in a single market firm, since the latter already has low monitoring costs, by virtue of having its headquarters in the same market.

Empirically, we measure the marginal effect of internal agglomeration in single-market firms by interacting *Firm-MSA establishments* with a new variable, *All hotels in one MSA*, which is equal to one if the establishment is in a firm with all of its hotels in the same MSA in a particular year, and is zero otherwise. The results are presented in Table 8 Column 1. The main effect of operating all of a firm's hotels in the same MSA is positive, suggesting that geographically

focused firms are relatively advantaged, on average. Meanwhile, the coefficient on *Firm-MSA establishments* remains positive, suggesting that different monitoring regimes do not explain the productivity effect from internal agglomeration. Interestingly, the coefficient on the interaction term is also *positive*, whereas we would have expected a negative coefficient, if monitoring and internal agglomeration effects were complementary in our setting. The test does not exhaustively evaluate all potential pathways improved monitoring might take, or even suggest that monitoring effects are unimportant; nevertheless, it is noteworthy that we do not find strong evidence of marginal monitoring effects.

The second mechanism we explore in depth is knowledge transfer. To do this, we first lay out three logical antecedents and then describe how these can be used to design a suggestive test of the knowledge transfer mechanism. First, if new establishments reflect ideas back onto pre-existing establishments, then the marginal effect of internal agglomeration should be greater when new establishments have more unique ideas. Second, if new establishments are more likely to scan the market for new ideas, because they are less rigid and more open to new ideas (Amburgey, Kelly, & Barnett, 1993; de Figueiredo, Rawley, & Rider, 2015), or because managers are actively engaged in developing the routines and processes that will be used to manage the establishment going forward, then new establishments should systematically be exposed to more ideas in the market place compared to pre-existing establishments. Third, if the originality of new ideas is correlated with changes in the productivity of the marketplace, then new establishments should be exposed to relatively more unique ideas when they operate in a more productive market. Given these three antecedents, we should expect that internal agglomeration effects will be magnified when internal agglomeration takes place in markets where productivity is increasing the fastest.

To evaluate the reflection effect empirically, we interact *Firm-MSA establishments* with *MSA productivity*. These results are presented in Table 8, Column 2. While the main effect of *Firm-MSA establishments* remains positive 0.003 (t-statistic = 2.9), the coefficient on the key interaction term is 0.039 (t-statistic = 7.5), meaning that increasing market-level productivity by one standard deviation (0.10) more than doubles the internal agglomeration effect, even after controlling for the main effect of market-level productivity and a full set of MSA × year fixed effects, which control for time-varying market-level heterogeneity. The interpretation is that increasing the productivity of rival establishments raises the productivity impact of internal agglomeration, a result that is consistent with the idea that new establishments reflect knowledge back onto pre-existing establishments.

We provide several additional checks on the finding that internal agglomeration causes productivity improvements. First, we conduct a placebo test in which we run the experiment as if it were lagged one period in the past. To implement the test, we run a difference-in-differences estimator taking lags of the dependent variable and all the right-hand side variables, except the key explanatory variables. If the placebo tests generated similar coefficient estimates on internal agglomeration, we would be concerned that our main results were spurious. In particular, such results would suggest that time-varying establishment specific characteristics unrelated to internal agglomeration were the factors driving the observed productivity effect. Instead, as we show in Table 8, Columns 3 and 4, we find that the placebo tests return coefficient estimates that are very close to zero for both *Firm-MSA establishments* (Column 3) and *positive treatment* and *negative treatment* (Column 4). The results give us some comfort that our main specification is correct.

A second set of robustness checks studies whether our internal agglomeration results, based on joint firm ownership of internally agglomerated establishments, are conflated with

franchising effects. The concern is that franchisors' efforts to coordinate and support franchisees would lead to productivity benefits, which might be conflated with the internal agglomeration effects we measure. While the existence of franchise effects need not invalidate internal agglomeration effects, and indeed would tend to bias our results toward zero, if the franchise "treatment" tended to be uncorrelated with the changes in internal agglomeration we measure, it is plausible that internal agglomeration and franchise effects could be conflated. For example, if establishments tended to become franchisees as they became more internally agglomerated—perhaps because small business owners tend to expand locally when they franchise one/some of their pre-existing establishments—it would be difficult to disentangle the two effects. Unfortunately, the Census did not systematically capture franchise information across the entire sample we study so we could not control for franchise effects directly in our main tests. However, as a robustness check, we use the franchisee and franchisor fields in the 2007 Economic Census to study how the inclusion of those variables influenced the measured internal agglomeration effect.

Column 5 of Table 8 runs a regression with establishment fixed effects that absorb the main effect of franchising, while controlling for the marginal effect of franchising, by allowing the

TABLE 9 Restaurants

Dep. variable = *productivity*	**(1)**	**(2)**
Firm-MSA establishments	0.001	
	(0.000)	
Positive treatment		0.035
		(0.001)
Negative treatment		0.017
		(0.001)
Log age	0.005	0.002
	(0.001)	(0.001)
Log MSA establishments	0.008	0.007
	(0.002)	(0.001)
Concentration index	−0.026	−0.026
	(0.007)	(0.006)
MSA productivity	0.121	0.104
	(0.012)	(0.010)
Log firm establishments ex-MSA	0.002	0.003
	(0.001)	(0.001)
Year fixed effects	Y	Y
Firm sales multiplier quartiles	Y	Y
Ownership change dummy	Y	Y
Establishment fixed effects	Y	Y
Year × MSA fixed effects	Y	Y
n	∼1,080,000	∼1,080,000
Adjusted R^2	0.70	0.70

Note: Standard errors are clustered at the establishment level.

franchisor and franchisee dummies to enter as interaction terms with the intensity of internal agglomeration term, *firm-MSA estabs*. In this case, the point estimate on the intensity of internal agglomeration term increases to 0.007 (t-statistic = 6.5), compared to 0.005 in Column 4 of Table 4 (our "baseline" result), providing further support for the claim that franchising effects do not bias the internal agglomeration results upward. Interestingly, the interactions between the intensity of internal agglomeration and the franchisee and franchisor dummies are negative, suggesting that internal agglomeration and franchise effects are substitutes on the margin.

4.4 | Results on restaurants

The hotel results suggest that internal agglomeration causes productivity to improve in a persistent manner, and that such effects appear to be driven, at least in part, by supply side effects such as knowledge transfer from new establishments to pre-existing establishments. However, one wonders whether such effects are idiosyncratic to the hotel context, or whether they generalize to other multi-unit establishment firms. To test whether the results apply out of context, we also apply the same tests to firms in the restaurant industry 1987–2007.

The restaurant results, summarized in Table 9, are quite similar to the hotel results, though there is some meaningful variation with respect to negative treatment effects. Table 9, Column 1, shows that increasing the intensity of internal agglomeration by one unit is associated with a precisely estimated 0.1% improvement in productivity. Decomposing the effect into persistent and contemporaneous effects using positive and negative treatment dummies, as we do in Column 2, reveals that the positive treatment effects are approximately 3.5%, about the same as for hotels. However, reducing internal agglomeration is also associated with an approximately 1.7% improvement in productivity, suggesting that restaurateurs tend to select restaurants for closure that are harming other establishments in the same firm, and that these effects swamp the effect of a loss of contemporaneous internal agglomeration. Broadly speaking the restaurant results suggest that the main results on the hotel sample generalize to other settings.

5 | CONCLUSION

Location choice is a crucial component of strategy for multi-unit firms, and internal agglomeration of sister establishments within the same firm in the same geographic area can lead to meaningful productivity effects. Indeed, internal agglomeration is a well-known phenomenon in many industries including auto manufacturing, retail, nursing homes, cable TV, and hospitality, to name a few. In this paper we study the magnitude and persistence (primarily with respect to state—internally agglomerated or not) of internal agglomeration effects by examining establishment-level productivity using microdata from the U.S. hotel industry from 1987 to 2007.

The results suggest internal agglomeration in the hotel industry causes meaningful within-establishment performance benefits: we find that doubling the intensity of internal agglomeration is associated with a productivity increase of about 2% in pre-existing establishments. The results are robust to alternate specifications and robustness checks, including the inclusion of MSA × Year fixed effects, matching on observables, and use of instrumental variables. Across specifications we find a similar positive relationship between internal agglomeration and productivity. Of course, increasing establishment productivity is not necessarily the same as

increasing firm profitability, as internal agglomeration also requires a, potentially large, initial investment, and may crowd out other investments, such that one should not conclude that internal agglomeration is "always good." Nevertheless, the results are striking, because they suggest that internal agglomeration does provide meaningful economic benefits to firms, on the margin.

We also find that persistent benefits (with respect to state) are at least as large as the contemporaneous benefits of internal agglomeration, though the latter have received far more attention in the literature on agglomeration. We replicate our main findings on Census microdata on the full population of U.S. restaurants from 1987 to 2007, suggesting that the internal agglomeration effects we document generalize to other industries with multi-unit firms.

We describe a range of mechanisms that could be driving the internal agglomeration effects we document, including demand and supply responses, monitoring effects, learning/knowledge spillovers, and resource sharing, and provide tests designed to disentangle the importance of each of these in our context. While tests of the marginal effect of the improved monitoring mechanism (Table 8, Column 1) and the resource sharing mechanism (Table 5, Columns 3 and 4) on the baseline internal agglomeration effect did not reveal strong evidence in favor of such mechanisms, there was suggestive evidence of intrafirm learning effects.

However, we acknowledge that our mechanism tests cannot rule out the possibility that other explanations play a role too. For example, it could be that as internal agglomeration increases, specialized labor in the region also increases, but when internal agglomeration decreases, specialized labor does not adjust very quickly (indeed, there is some evidence that U. S. labor mobility has decreased over time [Molloy, Smith, & Wozniak, 2011]). Under this scenario, the benefits from specialized labor would appear to be relatively persistent, and thus also consistent with our findings. While our interviews with industry experts suggested that sticky specialized labor was unlikely to be driving our results, we acknowledge that we cannot directly test this alternative explanation in the data. Other mechanisms that are unobservable to the econometrician may be at work too. For example, internal agglomeration may cause corporate managers to focus attention and effort on a location.[13]

Our findings have implications for both research and practice. For firms, this research suggests that internal agglomeration is important, and that a firm's internal agglomeration strategy may be enhanced by the extent to which it allows new establishments to explore and reflect new ideas back onto pre-existing establishments. For scholars, this paper represents a step toward unpacking internal agglomeration effects, allowing one to compare the relative importance of the drivers of internal agglomeration costs and benefits. Our findings may also have implications for the literature on performance effects of acquisitions (Mingo, 2013; Ramos & Shaver, 2013; Schoar, 2002). For example, our results suggest that acquisitions of nearby companies should benefit parent firms more so than acquisition of distant (but otherwise similar) companies. Questions remain about why contemporaneous effects vary substantially across industries and conditions under which positive contemporaneous effects (such as resource

[13]It is also worthwhile to note that the tendency of firms to seek out internal agglomeration benefits may vary over the industry lifecycle (Wang, Madhok, & Xiao Li, 2014). Similarly, it may be that new technologies change the relative costs and benefits associated with internal agglomeration over time. In our case, one might expect that the advent of the Internet would have made monitoring easier and lowered the cost of knowledge sharing between establishments owed by the same parent. We control for such temporal effects directly using year dummies. While future research may want to conduct a more in-depth study of the effect of the Internet on knowledge spillovers in the hotel or other industries, we did not find that our results differed significantly between "pre-Internet" (i.e., 1987–1997) and "post-Internet" (i.e., 2002–2007) time periods.

sharing) outweigh negative contemporaneous effects (such as cannibalization). For example, it would be interesting to explore complementarities between internal agglomeration and other firm characteristics, such as the relatedness of its operating units. More broadly, connecting contemporaneous and persistent internal agglomeration effects more directly to micro mechanisms appears to be a fascinating research agenda that this paper has only begun to address.

ACKNOWLEDGEMENTS

We are grateful to Janet Bercovitz, Kira Fabrizio, Ha Hoang, Jenny Kuan, Michael Leiblein, Susan Lu, Joanne Oxley, Maggie Zhou, Arvids Ziedonis, Rosemarie Ziedonis, seminar participants at BI Norwegian Business School, Copenhagen Business School, ESMT, NYU Stern School of Business, Ohio State University, the University of Connecticut, the University of Michigan, the University of Minnesota, Washington University, for valuable comments and suggestions. The research in this paper was conducted while the authors were Census Bureau research associates at the New York City Census Research Data Center (NYCRDC). Research results and conclusions expressed are our own and do not necessarily indicate concurrence by the Bureau of Census. This paper has been screened to insure that no confidential data are revealed. Financial support for this work was provided to Rawley by the Mack Institute at Wharton and to Seamans by the Ewing Marion Kauffman Foundation. The contents of this publication are solely the responsibility of the authors.

REFERENCES

Alcacer, J. (2006). Location choices across the value chain: How activity and capability influence agglomeration. *Management Science*, *52*(10), 1457–1471.

Alcacer, J., & Chung, W. (2014). Location strategies for agglomeration economies. *Strategic Management Journal*, *35*(12), 1749–1761.

Alcacer, J., & Delgado, M. (2016). Spatial organization of firms and location choices through the value chain. *Management Science*, *62*(11), 3213–3234.

Allen, T. J. (1984). *Managing the flow of technology: Technology transfer and the dissemination of technological information within the R&D organization*. Cambridge, MA: MIT Press Books.

Amburgey, T. L., Kelly, D., & Barnett, W. P. (1993). Resetting the clock: The dynamics of organizational change and failure. *Administrative Science Quarterly*, *38*(1), 51–73.

Appleyard, M. M. (1996). How does knowledge flow? Interfirm patterns in the semiconductor industry. *Strategic Management Journal*, *17*(2), 137–154.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, *68*(1), 29–51.

Baum, J. A. C., & Ingram, P. (1998). Survival-enhancing learning in the Manhattan Hotel Industry, 1898–1980. *Management Science*, *44*(7), 996–1016.

Bettis, R., Gambardella, A., Helfat, C., & Mitchell, W. (2014). Quantitative empirical analysis in strategic management. *Strategic Management Journal*, *35*(7), 949–953.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, *87*(1), 115–143.

Brandenburger, A., & Stuart, H. (2007). Biform games. *Management Science*, *53*(4), 537–549.

Bresman, H., Birkinshaw, J., & Nobel, R. (1999). Knowledge transfer in international acquisitions. *Journal of International Business Studies*, *30*, 439–462.

Bun, M. J. G., & Windmeijer, F. (2010). The weak instrument problem of the system GMM estimator in dynamic panel data models. *Econometric Journal*, *13*(1), 95–126.

Canina, L., Enz, C. A., & Harrison, J. (2005). Agglomeration effects and strategic orientations: Evidence from the US lodging industry. *Academy of Management Journal*, *48*(4), 565–581.

Chung, W., & Kalnins, A. (2001). Agglomeration effects and performance: A test of the Texas lodging industry. *Strategic Management Journal*, *22*(1), 969–988.

Darr, E. D., Argote, L., & Epple, D. (1995). The acquisition, transfer and depreciation of knowledge in service organizations: Productivity in franchises. *Management Science*, *41*(11), 1750–1762.

de Figueiredo, R. J. P., Rawley, E., & Rider, C. (2015). Why are firms rigid? A general framework and empirical tests. *Organization Science*, *26*(5), 1502–1519.

Fabrizio, K. R., & Thomas, L. G. (2012). The impact of local demand on innovation in a global industry. *Strategic Management Journal*, *33*(1), 42–64.

Feldman, M. P. (2000). Location and innovation: The new economic geography of innovation, spillovers, and agglomeration. *The Oxford Handbook of Economic Geography*, *1*, 373–395.

Folta, T. B., Cooper, A. C., & Baik, Y. S. (2006). Geographic cluster size and firm performance. *Journal of Business Venturing*, *21*(2), 217–242.

Gilbert, B. A., McDougall, P. P., & Audretsch, D. B. (2008). Clusters, knowledge spillovers and new venture performances: An empirical examination. *Journal of Business Venturing*, *23*(4), 405–422.

Giroud, X. (2013). Proximity and investment: Evidence from plant-level data. *The Quarterly Journal of Economics*, *128*(2), 861–915.

Glaeser, E. L., Kim, H., & Luca, M. (2017). Nowcasting the local economy: Using yelp data to measure economic activity. No. w24010. National Bureau of Economic Research

Greenstone, M., Hornbeck, R., & Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, *118*(3), 536–598.

Gupta, A. K., & Govindarajan, V. (2000). Knowledge flows within multinational corporations. *Strategic Management Journal*, *21*(4), 473–496.

Henderson, J. V., & Ono, Y. (2008). Where do manufacturing firms locate their headquarters? *Journal of Urban Economics*, *63*(2), 431–450.

Henderson, V. (2003). Marshall's scale economies. *Journal of Urban Economics*, *53*(1), 1–28.

Holtz-Eakin, D., Newey, W., & Rosen, H. S. (1988). Estimating vector autoregressions with panel data. *Econometrica*, *56*(6), 1371–1395.

Iacus, S. M., King, G., & Porro, G. (2011). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association*, *106*(493), 345–361.

Ingram, P., & Baum, J. A. C. (1997). Chain affiliation and the failure of Manhattan hotels, 1898–1980. *Administrative Science Quarterly*, *42*(1), 68–102.

Ingram, P., & Roberts, P. W. (2000). Friendships among competitors in the Sydney hotel industry. *American Journal of Sociology*, *106*(2), 387–423.

Jin, G. Z., & Leslie, P. (2009). Reputational incentives for restaurant hygiene. *American Economic Journal: Microeconomics*, *1*(1), 237–267.

Kalnins, A., & Lafontaine, F. (2013). Too far away? The effect of distance to headquarters on business establishment performance. *American Economic Journal: Microeconomics*, *5*(3), 157–179.

Krugman, P. (1991). *Geography and trade*. Cambridge, MA: MIT Press.

Lu, S. F., & Wedig, G. J. (2013). Clustering, agency costs and operating efficiency: Evidence from nursing home chains. *Management Science*, *59*(3), 677–694.

Marshall, A. (1920/1890). *Principles of economics*. London, England: Macmillan Publishers.

Martin, P., Mayer, T., & Mayneris, F. (2011). Spatial concentration and plant-level productivity in France. *Journal of Urban Economics*, *69*(2), 182–195.

McCann, B. T., & Folta, T. B. (2008). Location matters: Where we have been and where we might go in agglomeration research. *Journal of Management*, *34*(3), 532–565.

McCann, B. T., & Vroom, G. (2010). Pricing response to entry and agglomeration effects. *Strategic Management Journal*, *31*(3), 284–305.

Mingo, S. (2013). The impact of acquisitions on the performance of existing organizational units in the acquiring firm: The case of an agribusiness company. *Management Science*, *59*(12), 2687–2701.

Molloy, R., Smith, C. L., & Wozniak, A. (2011). Internal migration in the United States. *Journal of Economic Perspectives*, *25*(3), 173–196.

Mowery, D. C., Oxley, J. E., & Silverman, B. S. (1996). Strategic alliances and interfirm knowledge transfer. *Strategic Management Journal*, *17*(2), 77–91.

O'Neill, J. W., Hanson, B., & Mattila, A. S. (2008). The relationship of sales and marketing expenses to hotel performance in the United States. *Cornell Hospitality Quarterly*, *49*(4), 355–363.

Pancras, J., Sriram, S., & Kumar, V. (2012). Empirical investigation of retail expansion and cannibalization in a dynamic environment. *Management Science*, *58*(11), 2001–2018.

Ramos, M. A., & Shaver, J. M. (2013). When individual locations affect the choice of multi-location acquisition targets. *Strategic Organization*, *11*(2), 125–155.

Rawley, E. (2010). Diversification, coordination costs, and organizational rigidity: Evidence from microdata. *Strategic Management Journal*, *31*(8), 873–891.

Rosenthal, S. S., & Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In J. V. Henderson & J.-F. Thisse (Eds.), *Handbook of regional and urban economics* (Vol. 4, pp. 2119–2171). Amsterdam: Elsevier, Chap. 49.

Sakhartov, A. V. (2017). Economies of scope, resource relatedness, and the dynamics of corporate diversification. *Strategic Management Journal*, *38*(11), 2168–2188.

Saxenian, A. L. (1994). *Regional advantage: Culture and competition in silicon valley and route 128*. Cambridge, MA: Harvard University Press.

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

Shaver, J. M., & Flyer, F. (2000). Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States. *Strategic Management Journal*, *21*(12), 1175–1194.

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, *65*(3), 557–586.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, *49*(2), 326–365.

Wang, L., Madhok, A., & Xiao Li, S. (2014). Agglomeration and clustering over the industry life cycle: Toward a dynamic model of geographic concentration. *Strategic Management Journal*, *35*(7), 995–1012.

Williamson, O. E. (2005). The economics of governance. *American Economic Review*, *95*(2), 1–18.

Woo, H. S., Cannella, A., Jr., & Mesquita, L. (2019). How intra-and interfirm agglomeration affect new-unit geographic distance decisions of multiunit firms. *Strategic Management Journal*, *40*(11), 1757–1790.

Zhou, Y. M. (2011). Synergy, coordination costs, and diversification choices. *Strategic Management Journal*, *32*(6), 624–639.

How to cite this article: Rawley E, Seamans R. Internal agglomeration and productivity: Evidence from microdata. *Strat. Mgmt. J.* 2020;1–29. https://doi.org/10.1002/smj.3200