Paper to be presented at
the DRUID16 20th Anniversary Conference
Copenhagen, June 13-15, 2016

Intra-Firm Spillovers? The Stock and Flow Effects of Collocation

Evan Rawley
Columbia University
Management
erawley@columbia.edu

Robert Seamans
New York University Stern School of Business
Management
rseamans@stern.nyu.edu

Abstract

We study how intra-firm collocation—geographic clustering of business establishments owned by the same parent company—influences performance, decomposing the collocation effect into stocks and flows to learn about the mechanisms behind intra-firm agglomeration. Using Census micro data on the full population of U.S. hotels and restaurants from 1977-2007, we find that doubling the intensity of intra-firm collocation is associated with a productivity increase of about 2%. Further analyses reveal that a significant component of the productivity gains are attributable to stock effects, in the sense that productivity effects persist after an establishment ceases to be collocated. The results are consistent with the idea that proximity to other establishments owned by the same parent firm facilitates knowledge transfer, which has broad implications for firm strategy.

Jelcodes:L80,Z00

Intra-Firm Spillovers? The Stock and Flow Effects of Collocation

Abstract

We study how intra-firm collocation—geographic clustering of business establishments owned by the same parent company—influences performance, decomposing the collocation effect into stocks and flows to learn about the mechanisms behind intra-firm agglomeration. Using Census micro data on the full population of U.S. hotels and restaurants from 1977-2007, we find that doubling the intensity of intra-firm collocation is associated with a productivity increase of about 2%. Further analyses reveal that a significant component of the productivity gains are attributable to stock effects, in the sense that productivity effects persist after an establishment ceases to be collocated. The results are consistent with the idea that proximity to other establishments owned by the same parent firm facilitates knowledge transfer, which has broad implications for firm strategy.

1. INTRODUCTION

Locational strategy is crucial for multi-unit firms. Intra-firm collocation—clustering new establishments in close geographical proximity to the firm's pre-existing establishments—creates frictions (e.g., cannibalization and internal politicking), but can also improve efficiency (e.g., through resource and knowledge sharing). And indeed, intra-firm clustering is a well-known phenomenon in many industries including auto manufacturing, retail, nursing homes, cable TV, and hospitality, to name a few. While scholars and practitioners tend to emphasize the intra-firm flow benefits of collocation due to local economies of scale, such as joint procurement (Ingram and Baum, 1997), and monitoring (Lu and Wedig, 2013), evidence that collocation improves intra-firm learning (Darr, Argote and Epple, 1995) suggests that stock benefits of collocation (e.g., knowledge transfer) may be of great importance as well. For example, four of Pfizer's nine R&D labs are located in New England, perhaps because proximity facilitates internal agglomeration effects (Alcacer and Delgado, 2015). At least two unaddressed, but interesting questions follow from this premise. First, how large and persistent are intra-firm collocation effects; do they matter, and how do they compare to flow effects? And, second, do intra-firm collocation benefits accrue to the firm's new establishments, existing establishments, or both? Intuitively one might expect pre-existing establishments to teach the new, but do new establishments help pre-existing ones learn too? This paper addresses the persistence question by studying how collocation influences productivity, using longitudinal variation in the incidence of within-establishment collocation to

decompose collocation effects into stocks and flows. To address the direction question—the question of which establishments benefit more—we separately measure the effects of within-firm collocation on the firm's new and existing establishments.

The empirical analyses examine the magnitude, persistence and direction of collocation effects using data from two key industries where firms' geographic strategies are important and heterogeneous—the U.S. accommodation ("hotel") and food service ("restaurant") industries1977-2007. Our baseline results show that within-firm collocation is associated with a 1.1-2.4% improvement in the productivity of new and existing establishments.[1] Further analyses show that the stock benefits are economically and statistically meaningful in both the hotel and restaurant industries, while net flow effects vary by context. The results suggest that the persistent stock benefits of within-firm collocation, such as knowledge transference, are at least as large as (net) flow effects, which has broad implications for spatial strategy, particularly in multi-unit firms.

The paper makes three key contributions. First, whereas much of the literature on location-based strategy studies inter-firm collocation effects, we document a positive relationship between intra-firm collocation and performance in two large-scale datasets on economically important industries. Interestingly, the robustness of the results to controls for selection effects is consistent with the idea that there are intra-firm "spillovers"—or causal intra-firm agglomeration effects—between collocated establishments. Second, we demonstrate conceptually and empirically how entry and exit can be used to disentangle the stock and flow effects of collocation. Comparing these intra-firm results to inter-firm agglomeration effects identified in other research reveals a surprising result: while knowledge spillovers are less important for inter-firm agglomeration (Alcacer and Chung, 2014), they are economically and statistically meaningful in the intra-firm context. Third, the results suggest that collocation facilitates the transfer of local knowledge across establishments in two ways: (i) new establishments learn from pre-existing establishments' experience and (ii) pre-existing establishments learn from new establishments'

[1] To put these results in context, Giroud (2013) finds that decreases in monitoring costs of establishments leads to productivity increases of 1.3-1.4% Kalnins (2006) reports that over a ten year period, hotels averaged a return on sales of 13.4%. (Many sources put restaurant industry return on sales in a similar range).

experiments. While the former result is intuitive (Banaszak-Holl, et. al., 2006), the latter speaks to the important, but understudied, role of managerial experimentation and reflection in facilitating intra-firm learning.

Our analyses are particularly compelling for several reasons. First, we have unusually broad and rich establishment-level data on a complete census of every hotel and restaurant ever operated (for at least five years) in the United States 1977-2007. Second, we offer relatively precise estimates of the magnitude, persistence, costs and benefits of intra-firm collocation. Specifically, we employ a differences-in-differences estimator with establishment and time fixed effects and market-specific time trends to capture the within-establishment changes in productivity levels associated with collocation. Third, the results are robust to a wide-range of alternative explanations including endogenous sorting based on pre-trends, market-level characteristics, or firm or establishment-specific differences. While our research design does not allow us to conclusively rule out all alternative explanations, we triangulate our quantitative findings with institutional knowledge gleaned from interviews with industry experts to provide deeper insights into intra-firm collocation effects.

2. THEORY AND LITERATURE

Agglomeration effects can be persistent (stocks) or fleeting (flows), and the effects can be manifest within a firm or between firms. Thus, there are four types of agglomeration effects: inter-firm flows, inter-firm stocks, intra-firm flows, and intra-firm stocks (please see Figure 1). These spillovers can be further categorized as benefits or costs to the focal firm or establishment.[2]

Inter-firm flows (quadrant I) include "classic" agglomeration spillovers (e.g., Marshall 1920/1890)—contemporaneous inter-firm benefits due to collocation. Examples of such spillovers include complementary demand whereby increased clustering drives customer traffic (Chung and Kalnins, 2001; McCann and Vroom, 2010), and complementary supply whereby increased clustering attracts specialized suppliers of resources (e.g., Saxenian, 1994). The costs associated with inter-firm flow effects can be

[2] It is useful to point out that we separate flow and stock effects for theoretical and analytical tractability. In practice it may however be difficult to separate "flows" from "stocks" in some cases. It may also be that flows can accumulate over time to become stocks (Dierickx and Cool, 1989).

thought of as negative spillovers from increased competition for customers or resources (e.g., Alcacer, 2006), leading either to lower prices, as in standard Bertrand competition models, or increased costs due to a reduction in bargaining power with suppliers (e.g., Brandenburger and Stuart, 2007).

Intra-firm flows (quadrant II) are not classical spillovers from one firm to another, but rather spillovers (or "spill–ins") that accrue to sister establishments owned by the same parent firm in close proximity to each other. Benefits include increased economies of scale in purchasing and branding (Jin and Leslie, 2009), enhanced monitoring (Lu and Wedig, 2013; Kalnins and LaFontaine, 2014), and ongoing learning and benchmarking (Baum and Ingram, 1998; Kalnins and Mayer, 2004). Costs include cannibalization in that different business establishments owned by the same parent firm are now competing for the same customers (Kalnins, 2004; Pancras, Sriram, and Kumar, 2012). While the extent of cannibalization can vary across firms (Bennett, Seamans and Zhu, 2015), an open question is the extent to which cannibalization costs are borne by the new establishment, existing establishment or both.

Inter-firm stock transfers (quadrant III) will typically involve knowledge spillovers, for example leakage of a process or production technique or technological capability (Appleyard, 1996; Mowery, Oxley, Silverman, 1996), but in principle could involve human as well as intellectual capital (Allen, 1984). A firm with valuable knowledge gains a competitive advantage over its rivals, but can lose this advantage if the knowledge leaks into the pool of localized knowledge accessible by its rivals (e.g., through employee mobility). An open question in this literature is the extent to which incumbents and entrants can equally benefit from these pools of knowledge or whether the effects are asymmetric in that entrants and incumbents benefit differentially from each other (Eeckhout and Jovanovic, 2002; Knott, Posen and Wu, 2009; Posen and Chen, 2013). Of course, given the potentially large gains from such knowledge spillovers, firms engage in costly mechanisms to imitate, if they are a lagging firm, or to deter imitation, if they are a leading firm.

One focus of this paper is on inter-firm stock transfers (quadrant IV), which occur when a sister establishment transfers valuable knowledge to other establishments owned by the same parent firm in the

same geographic area. Our emphasis on localized knowledge transfer, 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 multi-national corporations transmit knowledge from business establishments in one country to another (e.g., Bresman, Birkinshaw and Nobel, 1999; Gupta and Govindarajan, 2000). Potential intra-firm collocation costs include coordination, bureaucratization and internal politicking (Williamson, 2005; Rawley, 2010; Zhou, 2011).

Figure 1 illustrates how our analysis maps to our contributions. Our first contribution is to document that intra-firm effects arising from collocation lead to performance improvements (i.e., we sum quadrants II and IV). These performance improvements are net of any inter-firm spillovers (i.e., quadrants I and III), which we control for in our regressions. Second, to the best of our knowledge, we are the first to study the relative contribution of stock and flow effects of intra-firm collocation on performance (i.e., we compare quadrants II and IV). The technique we describe can be used to disentangle stock and flow effects of intra-firm collocation in other contexts, and could even be used to disentangle stock and flow effects of inter-firm collocations (i.e., one could compare quadrants I and III). Third, we compare the effect of intra-firm collocation on performance for incumbent establishments and new establishments (i.e., we sum quadrants II and IV separately for both types of establishments). While the existing literature has focused on these differences in inter-firm settings, we emphasize intra-firm effects, showing that collocated incumbent and new establishments receive similar benefits from collocation. Interviews with industry experts qualitatively help to further elucidate the sources of these effects in our particular context.

3. RECOVERING STOCK AND FLOW EFFECTS

Before turning to our empirical tests we show analytically how one can exploit positive and negative collocation events to separate stock and flow effects (i.e., separately measuring the effects in quadrants II and IV in Figure 1). Suppose an establishment's performance π is a function of its initial endowment π_0 plus the benefits B and costs, $C>0$, of collocation, and a mean zero random disturbance term e. If the costs and benefits of collocation arrive as persistent stock effects, $\bar{u}>0$ and $\bar{a}>0$, respectively, and as flow

costs and benefits, c>0, b>0, respectively, only when the establishment is collocated, and zero otherwise, then we have:

$$(1) \qquad \pi_t = \pi_0 + B_t - C_t + e_t = \pi_0 + \bar{a} - \bar{u} + b_t - c_t + e_t.^3$$

Therefore, if an establishment becomes collocated at time t due to the arrival of another establishment in the same market owned by the same firm (which we will call a "sister" establishment below), the expected change in performance is:

$$(2) \qquad E(\Delta^+\pi) = \bar{a} - \bar{u} + b_t - c_t,$$

where there are net benefits to collocation when $\bar{a} + b_t > \bar{u} + c_t$. By contrast, if the collocated sister establishment subsequently exits, say at time t+1, the expected change in performance is:

$$(3) \qquad E(\Delta^-\pi) = b_{t+1} - b_t - (c_{t+1} - c_t) = c_t - b_t,$$

where $b_{t+1} = c_{t+1} = 0$ by definition, there is a reduction in the flow benefits due to collocation when $c_t < b_t$.

Taking expressions (2) and (3) together we can recover the net stock benefit of collocation by subtracting the change in performance from exits of local sister establishments from the change in performance from entry of local sister establishments:

$$(4) \qquad E(\Delta^+\pi) - E(\Delta^-\pi) = \bar{a} - \bar{u}.$$

Below, we use the empirical analog of (4) to estimate the net stock benefits of collocation, and therefore establish a lower bound on the magnitude and direction of positive stock transfers when establishments collocate, while using the empirical analog of (3) to establish the net flow benefits of resource sharing due to collocation. Note that in cases where cannibalization or other flow costs are relatively high, such as in the restaurant industry (Pancras, Sriram, and Kumar, 2012; Quelch, 2008), we are likely to find negative flow effects; that is, $b_t < c_t$. By comparing the results of our estimates of expressions (3) and (4), we can recover the relative contribution of (net) stocks and flows to intra-firm spillovers arising from collocation.

4. DATA AND EMPIRICAL STRATEGY

4.1. Data and setting

[3] Given expression (1), changes in performance for the establishment due to a change in its collocation status can be written as changes in the costs and benefits of collocation: $\Delta\pi = \Delta\bar{a} - \Delta\bar{u} + \Delta b - \Delta C + \Delta e$.

The data come from the U.S. hotel and restaurant industries from the 1977-2007 quinquennial Economic Census of Services and Economic Census of Retail Trade, $100 billion (Kalnins, 2006) and $200 billion (Basham and Menza, 2007) industries by revenue, respectively. The data includes the revenue and payroll (but not capital expenditures) of every hotel and restaurant in the United States, tracked longitudinally with a unique identifier, along with location and firm (but not brand) identifiers. 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 along with other unusual observations—establishments that are in the smallest 10% of the size distribution (by revenue) and observations in the 1st and 99th percentile of the labor factor productivity distribution.[4]

A number of potential stock and flow benefits associated with collocation have been widely reported in the scholarly literature on organizations including joint procurement, and more efficient centralized monitoring. These benefits were clearly echoed in our interviews with industry executives, analysts, consultants and managers. Interestingly, knowledge-based benefits of collocation were also noted to be important. 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 the cleaning cart to speed changeovers—spread across collocated 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. He explained that new establishments are often managed by seasoned managers from pre-existing hotels in the same market. Because some customization is expected

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

at a new establishment, it is common for the manager to be given some latitude to try out new ideas in the new setting. Through formal and informal means these managers will often apprise their former colleagues in the pre-existing hotel about the success or failure of such "experiments." Thus, the experience of industry experts support the idea that stock knowledge transfers may be an important source of intra-firm agglomeration effects. Below we take these anecdotal observations to the data to see if we can find evidence that such effects are systematic.

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 pick up the knowledge-based "stock" collocation effects that we are particularly interested in, while small enough to meaningfully capture a local area for the purposes of defining entry and exit. Other studies of collocation in the hotel and restaurant industries have used other, typically smaller, market definitions, particularly when the studies focused on micro-market "flow" collocation effects, such as referrals (Ingram and Roberts, 2000) or physical asset sharing. Our MSA-based market definition capture such flow effects, but also have the advantage of creating markets that are more geographically isolated from one another (i.e., compared to smaller market definitions). Furthermore, we verify that all of the key results hold using counties, which are substantially smaller than MSAs, as our market definition.

For the analyses described below, we use four cuts of the data: two from the hotel industry and two from the restaurant industry. For our main sets of results on incumbent performance, we focus on within-establishment changes in productivity from collocation. Observations that appear only once in the panel, those outside of metropolitan statistical areas, and establishments that are local monopolies are excluded, though the results are not sensitive to including these observations. The resulting dataset used for the main set of hotel results contains 58,648 hotels and 180,583 hotel-years from 23,985 firms. The resulting dataset used for the main set of restaurant results contains 313,609 restaurants and 856,599 restaurant-years from 146,416 firms. Summary statistics on these datasets, which we refer to as the "incumbent datasets" below, are presented in Table 1.

An additional set of results on entrant performance is generated by comparing the productivity of new establishments that are collocated with an incumbent to new establishments owned by the same firm that are not collocated. We use observations from 1977 to establish a baseline set of incumbents in a market. We then create a dataset of establishments that enter at any point from 1982 – 2007. These establishments remain in the dataset for all subsequent years until 2007 or the establishment exits, whichever comes first. As with the incumbent dataset, we exclude establishments located outside of metropolitan statistical areas (MSA), and local monopolies**.** The resulting datasets contains 261,906 hotel-years and 1,459,118 restaurant-years. Summary statistics for these datasets, which we refer to as the "entrant datasets" below, are presented in Table 2.

4.2. Effect of collocation on incumbent establishments

In our baseline tests we estimate the impact of changes in the intensity of collocation on establishment-level changes in revenue using a differences-in-differences estimator on the full panel. Specifically, for establishment i in year t we regress log revenue, y, on log payroll, l, 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 **X** indexed by c:

$$(5) \qquad y_{it} = \alpha + \beta_l l_{it} + \beta_1 ESTABS_{it} + \lambda_i + \mathbf{T_t} + \mathbf{X_{cit}B_c} + \varepsilon_{it}.$$

$ESTABS_{it}$ is the log count of the number of establishments owned by firm j in market g (MSA) over time, where establishment i is in firm j, ($ESTABS_{it}$= log $establishments_{it}$= log $establishments_{jgt}$).

The vector **X** contains variables that proxy for other potential shifters of an establishment's productivity frontier, including: log establishment age, measured as 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; the average productivity of other establishments in the same market (MSA LFP_{-igt}); 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 the number of establishments owned by the same firm in other markets (plus one); a

firm productivity spline (quartiles), measured as the average productivity of all the firm's other establishments (i.e., excluding i); and in the most demanding specification, hundreds of market-year fixed effects. All dollar-denominated values are deflated to 1977 dollars using CPI deflators from the Bureau of Labor Statistics; (one 1977 dollar is worth about four dollars today). ε_{it} is a mean-zero, normally distributed error term. Standard errors are clustered at the establishment-level.

Notably amongst the controls, log count of competitors' establishments in the same market (log establishments$_{gt}$), both controls for inter-firm agglomeration effects, and delivers a useful benchmark against which to judge the magnitude of any intra-firm effects. Market-year fixed effects are also particularly important, as they control for all market-level variation that might be correlated with collocation decisions and productivity (e.g., supply and demand shocks), ruling out a large set of confounds.

Expression (5) decomposes residual revenue (i.e., revenue not explained by labor) into components, including an establishment-specific component, and into time-varying within-establishment components. Given the specification—a log-log one-step productivity regression with shift parameters—the coefficient on ESTABS$_{it}$ (β_1) can be interpreted as a productivity elasticity due to intra-firm collocation. Specifically, β_1 measures the within-establishment percentage change in residual revenue given a percentage change in the number of collocated establishments within the same firm, conditional on $\mathbf{X_c}$. In other words, it captures the responsiveness of establishment-level productivity to the intensity of firm collocation.

The data are rich, but do not include two variables of interest. First, brand information is not reported consistently to the Census Bureau so our estimate of the effect of collocation should be interpreted as firm-level, not brand-level, collocation effects. For example, 48% of Host Hotels's 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. The effects we estimate are at the level of the owner (e.g., Host Hotels) and not at the level of the brand (e.g., Sheraton). Since hotels and restaurants are often franchised, the absence of brand information means that there will be markets where two establishments

of the same brand (e.g., Hilton hotels or Taco Bell) are collocated but are owned by two different franchisees. To the extent that brand collocation effects are meaningful (e.g., as in Ingram and Baum 1997; Bernstein, 2015), our firm-level results will underestimate the "true" collocation effect in two ways. First, by confining our estimates to within-firm collocation we miss positive collocation effects that operate at the brand, but not the firm-level. Second, by comparing firm collocation "treatments" against control group establishments that benefit from brand collocation effects, we overestimate the baseline against which firm collocation effects are measured, and therefore underestimate firm collocation effects (i.e., because we have a differences-in-differences specification). While we would rather capture both firm and brand-level collocation effects, the former appears to be understudied in the literature. And yet, one might expect within-firm collocation effects to be quite important, even when a firm operates multiple brands.

Second, we do not observe capital. Variation in capital stock is controlled for directly with establishment-level fixed effects, however, if capital expenditures are systematically larger (smaller) at pre-existing establishments when the firm increases (decreases) the number of establishments it owns in the same market, our productivity estimates would be biased upward due to omitted variable bias. Fortunately, capital flows should be highly correlated with labor flows and several variables in $\mathbf{X_c}$, including: establishment age, secular trends, and local market by year fixed effects. Furthermore, while initial capital expenditures are substantial, capital 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.

4.3. Effect of collocation on entrant establishments

We also study the effect of collocation on productivity of the entering establishments. For establishment i in year t we regress log revenue, y, on log payroll, l, a collocated entry dummy, $COLO_i$, with firm, δ, and year, T, fixed effects and vector of controls $\mathbf{X_c}$:

$$(6) \qquad y_{it} = \alpha + \beta_l l_{it} + \beta_{COLO} COLO_i + \delta_j + \mathbf{T_t} + \mathbf{X_{cit}B_c} + \eta_{it}.$$

In specification (6) the key explanatory variable COLO is an indicator equal to one if establishment i, part of firm j, enters a market g (MSA) where j already owns an existing establishment. Thus, the interpretation of β_{COLO} is: the average (over time) productivity level of entering a market as a collocated establishment relative to the productivity of other establishments from the same firm that enter markets where they are not collocated. The vector **X** is as in (5) above, and η_{it} is a mean-zero, normally distributed error term. Standard errors are clustered at the establishment-level.

4.4. Learning and resource sharing effects of collocation

Specification (5) is useful for estimating the correlation between collocation and performance, but is subject to at least two important limitations. First, it does not allow one to separately identify the stock (Figure 1, quadrant IV) and flow (Figure 1, quadrant II) effects of collocation because β_1 does not distinguish between positive and negative changes in the intensity of collocation. When $\beta_1 > 0$ all we can infer is that within-firm performance improves with (or is positively correlated with) intra-firm collocation.

Second, one can only interpret the coefficient on $ESTABS_{it}$ as the causal effect of collocation on productivity if there are no omitted variables in specification (5) that are correlated with the dependent variable (log revenue) and the error term. In other words, the identifying assumption is there are no unobservable factors missing in specification (5) that would both increase (decrease) revenue for the focal establishment and increase (decrease) the establishment's intensity of collocation. While it seems plausible that β_1 is well identified in specification (5)—it controls for time-invariant establishment-specific factors, market level trends, and proxies for capital flows—it is a treatment on the treated (TOT) estimator, which does not control for heterogeneity in the selection process by which firms choose markets in which to add a new establishment or eliminate an existing one. If the decision to collocate is correlated with time-varying establishment-specific trends, which are uncorrelated with market-level trends, then the TOT estimator will be biased relative to the "true" average treatment effect.

We deal with the key measurement and selection issues in turn. First, to disentangle the stock and flow effects, (and to more closely map the empirics to the discrete measure of collocation in the analytical model), we discretize changes in collocation intensity into positive and negative categorical variables and study the asymmetries between these two effects, using:

$$(7) \qquad y_{it} = \alpha + \beta_l l_{it} + \beta_{pos} COLO+_{it} + \beta_{neg} COLO-_{it} + \lambda_i + \mathbf{T_t} + \mathbf{X_{cit}B_c} + \varepsilon_{it},$$

where COLO+ and COLO- are categorical variables that are equal to one in the period in which the firm collocates (+) or ceases to collocate (-) in a particular market, and are zero otherwise. β_{pos} captures positive collocation effects, including both (net) stock and flow effects. Negative collocation effects, captured by β_{neg}, however, only pick up a reduction in the net flow effects of collocation, as any stock effects, 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 collocation we can separately recover stock and flow effects of collocation on productivity: flow benefits of collocation = $-\beta_{neg}$, stock benefits of collocation = $\beta_{pos} + \beta_{neg}$.

Second, to deal with the potential for selection bias in the firm's collocation decision, we use propensity score matching and Coarsened Exact Matching (CEM) to identify control group establishments that are similar along all observable dimensions to establishments where collocation status changes. The idea is to identify control group establishments that "look like" positively and negatively treated establishments prior to the treatment to disentangle treatment effects from pre-trends. Specifically, with propensity score matching we specify a logit predicting a positive change in collocation, and a logit predicting a negative change in collocation:

$$(8) \qquad COLO+_{it} = f(l_{it-1}, \mathbf{T_t}, \mathbf{X_{cit-1}}),$$

$$(9) \qquad COLO-_{it} = f(l_{it-1}, T_t, \mathbf{X_{cit-1}}),$$

where $\mathbf{X_{cit-1}}$ includes all of the controls from (1) and (2) lagged one period (i.e. 5 years, given the nature of the Census data) except the market-year fixed effects. The logits establish the propensity score for existing establishments of the probability of receiving a "treatment," where the propensity score

summarizes all of the available observable information about the selection process by which establishments are sorted into treatment and control groups. Using the common support of the propensity score distributions—the predicted values for treatment and control group observations from (3) and (4)—after trimming off outliers, we match non-treated establishments one-to-one to treated establishments.

We also match positive and negative treatments to two sets of control group observations using the Coarsened Exact Matching (CEM) technique developed by Iacus, King and Porro (2011). CEM is conceptually similar to propensity score matching in that it matches treated observations to "similar" control group observations based on all observable characteristics of both. However, instead of matching observations based on propensity scores, CEM matches exactly on a bundle of discretized characteristics, which often makes the matching procedure substantially more efficient computationally compared to propensity score matching. Another key difference between the two approaches is that CEM eliminates the need for balance checking; that is, because CEM matches exactly, one need not ensure that the mean of the distribution of each variable in the matching algorithm is similar after matching, as in propensity score matching. In other words, because each observation matches exactly along all observable dimensions we care much less about the exact distribution of the covariates when using CEM compared to when using propensity score matching. While eliminating balance checking has certain advantages—namely it eliminates ad hoc decisions about where to cut outliers—it is not a panacea, as it is not necessarily conceptually obvious how much to coarsen the data without balance checking the result afterward to see if the variables have been coarsened "enough." Furthermore, intuitively one may have more confidence in a matching estimator when it results in ex ante balance between treatment and control groups, even if such balance is not required.

Once the two matched samples are identified (i.e., from (8) and (9)), they are combined and analyzed using specification (5). If any remaining unobservable establishment-specific time-varying information does not affect both the collocation selection process and changes in revenue at the focal establishments, then the matched sample will be equivalent to random assignment of collocation to legacy establishments.

That is, if the matched sample represents a valid counterfactual to the treatment group, one can interpret the coefficient on collocation intensity the average treatment effect of intra-firm collocation.

5. RESULTS

5.1. Summary statistics

The summary statistics for the incumbent and entrant datasets in Tables 1 and 2, respectively, show that collocation intensity varies by industry. Also, the raw data indicate that firms and establishments in our sample are relatively small—a typical firm in our study operates 6-8 establishments with total revenue of approximately $5-8 million ($20-30 million in 2015 dollars).[5] 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 or restauranteur, like the Marriott Corporation or McDonalds, but rather a small business owner of a handful of small to mid-sized establishments.

Table 1, (left panel), shows that churn influenced collocation at a relatively modest rate in the hotel sample: a lone hotel was joined by another hotel from the same firm in the same market in 2% of the observations (positive treatment), and a previously collocated hotel became the only hotel owned by its parent firm in the market (MSA) in 1% of the observations (negatively treated). The mean of ESTABS (= log $establishments_{it}$) is 1.64 in the incumbent hotel sample, which means that the average incumbent hotel is collocated with approximately five ($e^{1.64}$=5.2) sister hotels (i.e., other hotels owned by the same parent firm in the same MSA). Per Census rules we are not permitted to report minimum and maximum values for covariates, but we can note that the distribution of ESTABS is highly skewed to the right, with tens of thousands of hotel years having a value of zero. Average annual hotel sales were $906,000, while the coefficient on log $establishments_{j\text{-}gt}$ is 0.61, meaning that the average incumbent hotel has approximately two sister hotels located in other MSAs ($e^{0.61}$=1.8).

[5] The mean number of establishments operated by a firm in the hotel sample was eight with sales of $906,000/establishment in 1993 dollars. Per the Bureau of Labor Statistics, 2015 dollars are worth 4.2 1993 dollars. Therefore, the average hotel firm in our sample had approximately $31M (=9 x $906,000 x 4.2) of revenue in 2015 dollars. In the restaurant sample the average firm operated 6.3 establishments with 1993 sales of $726,000, or $19M (=6.3 x $726,000 x 4.2) of revenue in 2015 dollars.

Churn had a much more substantial impact on collocation in the restaurant sample: a lone restaurant was joined by another restaurant from the same firm in the same market in 7% of the observations (positive treatment), and a previously collocated restaurant being the only restaurant owned by its parent firm in 6% of the restaurant-years (negative treatment). Yet, ownership in the hotel industry is more concentrated geographically than in the restaurant industry. In the restaurant incumbent sample (Table 1, right panel), the coefficient on ESTABS is 0.67—the average incumbent restaurant is collocated with approximately two other sister restaurants ($e^{0.67}$=2.0). Average annual restaurant sales were $726,000, while the coefficient on log establishments$_{j\text{-}gt}$ is 1.18, meaning that the average incumbent restaurant has approximately three sister restaurants located in other MSAs ($e^{1.18}$=3.3).

5.2. Baseline results for incumbents and entrants

We investigate the effect of within-establishment changes in collocation on establishment performance in Table 3, where we present the baseline results corresponding to specification (5) above. In Column 1, log sales is regressed against ESTABS, log payroll, and establishment and year fixed effects. The coefficient on ESTABS is 0.021, which means doubling the number of establishments owned by the same firm in the same market (MSA) is associated with a 2.1% increase in productivity, and is statistically significant at the 5% level. Although the regression is quite parsimonious, the adjusted R^2 is 0.91, because sales are highly correlated with labor.[6] The high R^2 suggests that any bias from omitted capital flows is likely to be small, though admittedly predicting changes in productivity is less precise than predicting changes in revenue.

The coefficient on ESTABS remains positive, significant, and of similar magnitude in column 2, even after including additional controls associated with the establishment, the firm, and the market. Inter-firm agglomeration effects (log establishments$_{gt}$) are indistinguishable from zero, and the point estimate is negative and much smaller in absolute value than the estimate on intra-firm agglomeration. The

[6] An alternative two-step productivity estimator yielded a very similar coefficient on ESTABS, but a much lower R^2 (i.e., in the second stage), reflecting the fact that heterogeneity in productivity is much harder to predict than heterogeneity in revenue.

coefficient on log age is positive and significant, even though aging is associated with declining productivity (Jones, 1999), probably because the variable is doing a good job picking up capital flows.

Columns 3 and 4 replicate the prior set of regressions on restaurants. The coefficient estimate on ESTABS in the restaurants sample is also positive and significant, though at 0.011 the point estimate is of lower absolute magnitude than in the hotels sample. In this setting the estimate of net inter-firm agglomeration effects are negative and statistically significant (-0.008): it appears that in the restaurant industry competition trumps positive inter-firm agglomeration effects. Taken together, the results from Table 3 provide suggestive evidence of a positive relationship between intra-firm collocation and performance.

Table 3 shows the relationship between collocation and productivity at pre-existing establishments. To compare the impact of collocation upon entry from new establishments to other new establishments that are not collocated when they enter, we present results on the "entrant" dataset using specification (6) from above in Table 4. In Column 1, log sales is regressed against a dummy indicating if the establishment was ever a collocated entrant, log payroll, and number of controls including year fixed effects. The coefficient on the collocated entrant dummy ($COLO_i$) is positive and significant (0.041), which suggests that new collocated establishments perform better than new establishments that are not collocated. To make the analysis more compelling, column 2 includes a firm fixed effect, which allows us to compare new collocated establishments against new solo establishments within the same firm. The coefficient on the collocated entrant dummy remains positive and significant, though the magnitude of the coefficient drops by close to 50% to 0.024, indicating a substantial amount of heterogeneity across firms in terms of the productivity of new establishments (i.e., better firms are more likely to collocate.)[7] The economic interpretation is that new collocated establishments are 2.4% more productive than new establishments from the same firm that are not collocated. Interestingly for new hotels, inter-firm agglomeration effects are positive and statistically significant, (though they small—on the order of 0.5-

[7] Alcacer (2006) finds that better firms in the cellular handset industry are less likely to collocate close to other firms. Our finding is with respect to the decision on the part of a given firm to collocate new establishments in close proximity with its existing establishments.

1%.) When contrasted with the within-establishment inter-firm results, these regressions suggest that inter-firm agglomeration effects are asymmetric for pre-existing versus new hotels, perhaps because new hotels are more receptive to learning from others.

Columns 3 and 4 replicate the prior set of regressions on restaurants. The coefficient on the collocated establishment dummy in the restaurant sample is also positive and significant, and of a similar magnitude (0.020 or 2%, in column 4) as in the hotel sample. For new restaurants inter-firm collocation effects are negative, just as for pre-existing restaurants, suggesting that, in this industry, MSA-level competition effects typically outweigh any positive inter-firm agglomeration effects.

While one cannot compare point estimates across the four samples with any statistical precision, the fact that the coefficient on collocation is similar between hotels and restaurants for new establishments (Columns 2 and 4 in Table 4), but greater for incumbent hotels than restaurants (Columns 2 and 4 of Table 3), suggests that either the benefits of intra-firm collocation are weakest for pre-existing restaurants, perhaps because they have the least to learn from new establishments, or, the costs of collocation are highest for incumbent restaurants, perhaps because newly collocated restaurants tend to cannibalize pre-existing sister establishments.

5.3. Disentangling stock and flow effects of collocation

While the positive association between collocation and productivity is suggestive of inter-firm spillovers due to collocation, we also want to better understand the drivers of this collocation effect. We, therefore, disentangle the stock and flow effects of collocation by discretizing changes in collocation intensity into positive and negative categorical variables, as in specification (7) above. In Table 5 Column 1 we study the hotel sample, where log sales is regressed against a positive treatment dummy, $COLO+_{it}$, and a negative treatment dummy, $COLO-_{it}$, which are equal to one in the year in which an existing establishment's collocation status changes positively or negatively, respectively, and are zero otherwise, as well as log payroll and a full set of controls including establishment and year fixed effects.

Column 2 also includes MSA-year fixed effects, which controls for market-specific time-trends in local supply and demand conditions that may jointly influence productivity and collocation. We find that positive treatment ($COLO+_{it}$) is a positively correlated with hotel productivity with negative treatment ($COLO-_{it}$) is negatively correlated with hotel productivity. Both are precisely estimated. The coefficient estimate of 0.033 on $COLO+_{it}$, and -0.014 on $COLO-_{it}$, which means that becoming collocated is associated with increasing productivity by 3.3%, while exiting from being collocated is associated with a 1.4% decline in productivity. The difference in magnitude between the two collocation dummies, a precisely estimated 1.8%, represents the net stock effect of intra-firm collocation; it is the persistent net benefit an establishment gains from being collocated. Inter-firm collocation effects are again negative and close to 1% in absolute value. The evidence on the hotel sample in Table 5 suggests that just over half of the benefit of intra-firm collocation is associated with stock benefits (quadrant IV in Figure 1), such as one-time learning effects, and just under half is associated with flow benefits (quadrant II).

We perform the same empirical tests on the restaurant sample in Table 5 columns 3 and 4, finding a precisely estimated positive treatment effect of collocation of 2.8% for restaurants, which is similar in direction and magnitude to the effect for hotels. However, we find the negative treatment effect of losing collocation is +1.1% for restaurants, which is the opposite sign than in the hotel sample. As with the hotel sample, comparing positive and negative treatment effects reveals that the stock effects of collocation are economically and statistically meaningful. The point estimates on inter-firm collocation effects are negative, but close to zero, (and even indistinguishable from zero in column 4).

The results suggest that restaurants benefit both when they become collocated and when they cease to be collocated. If positive and negative treatment effects were randomly assigned, this last result would be quite puzzling. How could restaurants benefit both from being collocated and from exiting from being collocated? One potential explanation is that there are positive net stock effects and negative net flow effects associated with collocation in the restaurant industry. Another possibility is that ex ante restaurant firms purposely open restaurants in the same market when they have the highest expected positive

collocation benefits, and ex post close collocated restaurants that experience the highest rates of cannibalization (and other flow costs). In other words, the endogeneity of the collocation decision potentially biases the coefficient estimates in our empirical tests (Shaver and Flyer 2000; Lu and Wedig, 2013). And, of course, even if the coefficient estimates in the hotel regressions fit more "naturally" with our intuition, the endogeneity critique may be applied to the hotel sample results as well.

5.4. Addressing threats to identification

While our baseline results suggest a positive relationship between intra-firm collocation and performance, even after controlling for time-invariant establishment specific characteristics, as well as a host of time-varying establishment, firm and market characteristics, it is still possible that our treatment on the treated estimator will be biased due to endogenous sorting. Clearly, firms choose where and whether to collocate strategically. The question is how large is the resulting bias. In the ideal experiment we would randomly assign collocation status across focal establishments. Unfortunately, it is not feasible to run such an experiment, and we could not identify any instruments that both met the exclusion restriction and were strong enough in the first stage to be worthwhile candidates for two-stage least squares estimates. Thus, to estimate average treatment effects net of selection on observables, we turn to a matching approach, as described in Section 4.4, as well as a battery of robustness checks.

Large ex ante differences between the treatment group and the (full) control group across the means of all the covariates suggest that matching is important, though also potentially challenging.[8] Still, we did find "good matches" for the hotel sample in the sense of being able to eliminate most statistical differences in the means of the observables between the treatment groups and their matched sample control groups, using propensity score matching. After eliminating observations that were off the common support of the joint distribution of the propensity scores for positive and negative treatment groups separately, and trimming off outliers at the 25th and 95th percentile of the joint propensity score distributions, the matching algorithm resulted in a substantial reduction in the difference in the means of

[8] Per Census Bureau disclosure restrictions we only describe the t-tests before and after matching in general terms.

all of the observable covariates. Whereas the t-tests on all of the means of the covariates were large and significant before matching, after matching, only log establishments$_{it}$ (5% level), and the MSA labor factor productivity variable (10% level) remained significantly different between the treatment and control groups; and even for these regressors the differences were much smaller. We obtained similar "second stage" results using Coarsened Exact Matching (CEM), though, as expected, the covariates did not balance as closely after matching as in the propensity score specification.

The fact that CEM does not deliver matched covariates as effectively as propensity score matching is not surprising given that CEM matches based on coarsened "buckets" of covariates and not a single predicted value (i.e., the propensity score). Ideally we would present both propensity score matched and CEM results anyway. However, because the Census frowns upon disclosing additional samples, we chose to present only the propensity score results. Intuitively it seems that results based on balanced covariates (i.e., the propensity score matched results) should be more reliable than results based on unbalanced covariates (i.e., the CEM results), even when treatment and control group observations are matched ex ante exactly in a coarse way. Thus, we only report the propensity score matched results for the hotel sample.

While matching was relatively "clean" for the hotel sample and the results quite consistent across the matched and "naïve" specifications, we had mixed results in restaurant sample. We could not obtain high quality matches using propensity score matching with the restaurant sample, in the sense that we could not find matches where most of the differences in the means of the covariates were indistinguishable from zero between the treatment and control group after matching. We, therefore, cannot claim that our second stage propensity score matched results are average treatment effects for the restaurant sample. Again, to avoid disclosing additional samples, we report only the CEM results in Table 6. While the CEM covariates also were not well balanced, the second stage estimates were stable across CEM specifications, and given that treatment and control group observations were matched "exactly", the CEM results are qualitatively superior to the unbalanced propensity-score matched results.

Table 6 presents the second stage of the matched sample results. The estimate on the continuous measure of collocation, ESTABS, in the matched sample in Column (1) is similar to the estimate obtained in the regression in Table 3 Column (2) for hotels.[9] The result is important because it suggests that selection on observables does not have a large impact on the magnitude of local within firm collocation effects in the hotel sample. Treated hotels are very different from most untreated hotels ex ante, but such heterogeneity does not appear to bias the results. Thus, in the absence of omitted variable bias, the net benefits of collocation can be interpreted as casual effects—intra-firm spillovers: random assignment of intra-firm collocation leads to performance benefits. Of course, the causal interpretation is based only on quasi-exogenous assignment of collocation status, and still allows for intra-firm benefits to be endogenously sought out once a firm chooses to collocate.

Column (2) performs the second stage matched sample analysis for the restaurants sample using the CEM matched sample. The coefficient on ESTABS is approximately 1.9% in the matched sample analysis compared to 1.1% in the unmatched sample. The interpretation is that while selection has a meaningful bearing on overall collocation effects in the case of restaurants, net collocation benefits are still positive, even controlling for selection on observables.

While the results are not as well identified as in a randomized control trial, the overall pattern of the evidence in both the hotel and restaurant samples is consistent with collocation exerting a positive causal effect on local within-firm performance. In other words, it appears that intra-firm spillovers are indeed economically and statistically significant.

5.5 Discussion

The results suggest that intra-firm collocation positively influences establishment-level productivity. We interpret this effect as evidence of intra-firm spillovers in the sense that randomly assigning collocation to an existing establishment leads to increased productivity. However, we do not wish to suggest that intra-firm spillovers are externalities, as with inter-firm spillovers. Indeed, once the intra-firm collocation

[9] Because the matched sample is only 7% of the overall hotel sample, a number of MSAs have only a single observation in each period. As such, we exclude the MSA x year fixed effects in this specification, though including them has no meaningful effect on the main result.

decision has been made, we expect firms to endogenously work hard to find the benefits of collocation and minimize the costs, whereas, with inter-firm collocation, firms will react to collocation by other firms by seeking ways to minimize the benefits that flow to others. Thus, as explicated by the four quadrants in Figure 1, intra-firm spillovers are conceptually related to, but still distinct from inter-firm spillovers. Both refer to the causal effects of collocation, but only the latter are externalities.

Relatedly, one might wonder why, given the benefits of collocation, all firms don't cluster their establishments in a single market. The most likely answer is that there are diminishing marginal returns to collocation—as with our log-log specification—along with some benefits of opening establishments in new markets, which operate outside of our regressions. For example, opening an establishment in a new market creates a new opportunity for collocation (i.e., in the new market) in the future.

The potential for unobservable changes in capital stock represents the most salient potential limitation of our empirical approach. If firms typically refurbish their existing establishments when they open a sister establishment in the local area, our analysis would conflate productivity effects with factor intensity effects. While our interviews with hotel executives suggested that such investment patterns would by no means be the norm—hotel firms generally reinvest in hotels according to age- and market-based considerations, not based on whether they collocate in those areas—we cannot completely rule out the potential for such effects to be at work. Our controls and matching approach do address this issue to a large extent. However, our empirical approach does have limitations. Because we can only match on the observable characteristics of the firms, markets and establishments in our sample, unobservable characteristics that are correlated with collocation and revenue still represent a potential endogeneity concern.

Another limitation of this research is that in some matching specifications the difference between the positive and negative treatment effects were noisy in the second stage (i.e., in the matched sample version of Table 5 where the results where positive and negative treatment enter separately). The point estimates of positive and negative treatment effects were always precisely estimated, and of the same size and

approximate magnitude as in Table 5, but the t-test on the sum of the COLO+ and COLO- coefficients was not distinguishable from zero at the five percent level in some matching specifications, perhaps because the sample size in the matched sample regressions were much smaller than in the full sample. Thus, we cannot rule out the possibility that stock effects related to collocation are driven in part by selection effects and are not intra-firm "spillovers" per se. While this ambiguity does not contradict the evidence in support of our main conceptual thrust, it would be better if we could statistically determine whether the magnitude of the stock effects remained after adjusting for selection effects. Thus, one direction for future research would be a better understanding of the extent to which firms endogenously collocate to realize intra-firm flow benefits (quadrant II) or intra-firm stock benefits (quadrant IV). As with inter-firm spillovers, we suspect there is substantial variation across industries (Appleyard, 1996).

Finally, another limitation of this research is that we can only offer qualified statements about the mechanisms by which intra-firm agglomeration effects obtain. From our interviews we expect that collocation leads to informal and formal knowledge sharing between proximate managers are the key driver of the stock effects we estimate.[10] However, we do not observe such effects directly. Rather we infer this kind of behavior from performance patterns in the data. As an extension of this work we explored potential mechanisms that might support the anecdotes about the costs and benefits of intra-firm collocation we heard from industry experts, by examining marginal effects that might give some additional insight into how intra-firm collocation relates to performance at a more micro level. Specifically, in untabulated results, we studied how the interaction between establishment age and collocation, and local competitors' productivity and collocation influenced productivity. The idea for studying these two interaction effects was that older establishments would be more rigid (de Figueiredo, Rawley, and Rider, 2015), and therefore less willing to learn from experiments by new establishments, while new establishments in high-productivity markets might capture, and reflect back, more valuable

[10] We also discussed alternative explanations for the stock effects we estimate in our interviews. For example, we felt intra-firm collocation might lead to one-time investments in long-lived shared assets (e.g., distribution centers for restaurants), but were told that while flow benefits from sharing suppliers and joint purchasing would be quite common, investments in shared fixed assets would be unusual in the hotel industry. Furthermore, most of the restaurant firms in our sample are too small to own their own regional distribution centers.

local knowledge. Both effects were statistically and economically significant and in the direction expected as interaction terms in the baseline regressions—the benefits of collocation were far less pronounced in older establishments, and more pronounced for establishments with high-productivity competitors—neither result was robust to matching, however.

6. CONCLUSION

Location choice is a central component of strategy for multi-unit firms. Collocation of different firms in the same geographic area can lead to inter-firm spillovers. Similarly, collocation of sister establishments in the same geographic area can lead to intra-firm spillovers. These intra-firm spillovers can have large effects on how establishments operate and perform. In this paper we study the persistence, magnitude and direction of intra-firm collocation effects by examining how collocation of sister establishments influences establishment level performance, and how stock effects of collocation, such as knowledge transfers, compare to recurring flow effects. We examine these issues using microdata from the U.S. hotel and restaurant industries—economically important industries where multiunit firms are prevalent—from 1977-2007.

The results suggest collocation causes economically and statistically meaningful positive within-firm performance benefits, for both new and incumbent establishments. Stock benefits are found to be at least as large as the flow benefits of collocation, though the latter have received far more attention in the literature on collocation. Our results are consistent with anecdotal evidence from hotel industry executives of how best practices diffuse locally within their organizations. Taken together, the evidence suggests that intra-firm collocation helps new establishments learn from existing sister establishments' experience, and helps existing establishments learn from new establishments' experiments. Existing establishments transfer location-specific knowledge from new establishments, and new establishments experiment with new approaches and techniques that can be transferred or "reflected" back to existing establishments.

The findings have implications for both research and practice. For firms, this research suggests that a firm's collocation strategy is enhanced by the extent to which it allows new establishments to experiment with new practices and techniques. For scholars, this paper represents a step toward unpacking collocation effects, allowing one to compare the relative importance of the drivers of collocation costs and benefits. Questions remain about why flow effects vary substantially across industries and conditions under which positive flow effects (such as resource sharing) outweigh negative flow effects (such as cannibalization). Moreover, connecting stock and flow effects to more directly to micro mechanisms appears to be a fascinating research agenda that this paper has only begun to address.

REFERENCES

Alcacer, Juan. 2006. Location Choices Across the Value Chain: How Activity and Capability Influence Collocation. Management Science 52(10): 1457-1471.

Alcacer, Juan, and Wilbur Chung. 2014. Location Strategies for Agglomeration Economies. Strategic Management Journal **35**(12): 1749-1761.

Alcacer, Juan, and Mercedes Delgado. 2015. Spatial organization of firms and location choice through the value chain. Management Science, forthcoming.

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.

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

Azoulay

Baum, Joel A.C., and Paul Ingram. 1988. Survival-enhancing Learning in the Manhattan Hotel Industry, 1898-1980. Management Science **44**(7): 996-1016.

Banaszak-Holl, Jan, Will Mitchell, Joel A.C. Baum, and Whitney B. Berta. 2006. Transfer learning in ongoing and newly acquired components of multiunit chains: U.S. nursing homes, 1991-1997. Industrial and Corporate Change **15**(1): 41-75.

Basham, Mark, and Justin Menza. 2007. Industry Surveys: Restaurants. Standard and Poors.

Bennett, Victor M., Robert Seamans and Feng Zhu. 2015. Cannibalization and option value effects of secondary markets: Evidence from the U.S. concert industry. Strategic Management Journal, forthcoming.

Bernstein, Shai, and Albert Sheen. 2015. The operational consequences of private equity buyouts: Evidence from the restaurant industry. Working paper.

Brandenburger, Adam, and Harborne Stuart. 2007. Biform games. Management Science, 53: 537–549.

Bresman, Henrik, Julian Birkinshaw, and Robert Nobel. 1999. Knowledge transfer in international acquisitions. Journal of International Business Studies: 439-462.

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

Darr, Eric D., Argote, Linda, and Dennis Epple. 1995. The acquisition, transfer and depreciation of knowledge in service organizations: Productivity in franchises. Management Science **41**(11): 1750-1762.

Dierickx, Ingemar, and Karel Cool. 1989. Asset stock accumulation and sustainability of competitive advantage. Management Science **35**(12): 1504-1511.

de Figueiredo, Rui J.P., Rawley, Evan and Christopher Rider. 2015. Why are firms rigid? A general framework and empirical tests. Organization Science, forthcoming.

Eeckhout, Jan, and Boyan Jovanovic. 2002. Knowledge spillovers and inequality. American Economic Review **92**(5):1290–1307.

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

Gupta, Anil K., and Vijay Govindarajan. 2000. Knowledge flows within multinational corporations. Strategic Management Journal, **21**(4), 473-496.

Iacus, Stefano M., Gary King, and Giuseppe Porro. 2011. Multivariate matching methods that are monotonic imbalance bounding. Journal of the American Statistical Association **106** (493): 345-361.

Ingram, Paul, and Joel A.C. Baum. 1997. Chain affiliation and the failure of Manhattan hotels, 1898-1980. Administrative Science Quarterly **42**(1): 68-102.

Ingram, Paul, and Peter W. Roberts. 2000. Friendships among competitors in the Sydney hotel industry. American Journal of Sociology **106**(2): 387-423.

Jin, Ginger Z., and Phillip Leslie. 2009. Reputational incentives for restaurant hygiene. American Economic Journal: Microeconomics, **1**(1), 237-267.

Jones, Peter. 1999. Hotel productivity: The McKinsey Global Institute study. Tourism and Hospitality Research, **1**(2): 178-181.

Kalnins, Arturs. 2004. An empirical analysis of territorial encroachment in franchised and company-owned branded chains. Marketing Science **23**(4): 476-489.

Kalnins, Arturs. 2006. The U.S. lodging industry. The Journal of Economic Perspectives **20**(4): 203-218.

Kalnins, Arturs, and Wilber Chung. 2004. Resource-seeking agglomeration: A study of market entry in the lodging industry. Strategic Management Journal **25**(7): 689-699.

Kalnins, Arturs, and Francine Lafontaine. 2014. Multi-unit ownership in franchising: Evidence from the fast food industry in Texas. RAND Journal of Economics **35**(4): 747-761.

Kalnins, Arturs, and Kyle J. Mayer. 2004. Relationships and hybrid contracts: An analysis of contract choice in information technology. Journal of Law, Economics, and Organization 20.1: 207-229.

Knott, Anne Marie, Posen, Hart E., and Brian Wu. 2009. Spillover asymmetry and why it matters. Management Science **55**(3):373–388.

Lu, Susan F., and Gerard J. Wedig. 2013. Clustering, agency costs and operating efficiency: Evidence from nursing home chains." Management Science **59**(3): 677-694.

Marshall, Alfred. 1920/1890. Principles of Economics. Macmillan Publishers: London, England

McCann, Brian T., and Govert Vroom. 2010. Pricing response to entry and agglomeration effects. Strategic Management Journal **31**(3): 284-305.

Mowery, David C., Oxley, Joanne E., and Brian S. Silverman. 1996. Strategic alliances and interfirm knowledge transfer. Strategic Management Journal **17** (2): 77-91.

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

Posen, Hart E. and John S. Chen. 2013. An advantage of newness: Vicarious learning despite limited absorptive capacity. Organization Science **24**(6): 1701-1716.

Quelch, John. 2008. Starbucks' lessons for premium brands. HBS Working Knowledge (July 9), http://hbswk.hbs.edu/item/5973.html.

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

Saxenian, AnnaLee. 1994. Regional Advantage: Culture and Competition in Silicon Valley and Route 128 Cambridge Massachusetts and London. Harvard University Press: Cambridge, MA.

Shaver, J. Myles, and Fredrick Flyer. 2000. Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States. Strategic Management Journal **21**(12): 1175-1194.

Williamson, Oliver E. 2005. The economics of governance. American Economic Review, 1-18.

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

Figure 1: Four types of agglomeration effects

	Inter-firm collocation	**Intra-firm collocation**
Flow	**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
Stock	**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

Table 1: Summary statistics for the incumbent datasets

	Hotel sample (n=180,583)		Restaurant sample (n=856,699)	
	Mean	Stdv	Mean	Stdv
Sales ($000)	906	4,913	725	940
Log establishments$_{it}$ (ESTABS$_{it}$)	1.64	1.35	0.67	1.16
Positive treatment (COLO+ $_{it}$)	0.02	0.15	0.07	0.25
Negative treatment (COLO-$_{it}$)	0.01	0.12	0.06	0.23
Payroll ($000)	254	1,445	203	307
Age (years)	6.0	6.5	6.5	6.8
Log establishments$_{gt}$	5.39	1.00	7.44	1.20
MSA LFP$_{-i,gt}$	0.00	0.10	-0.02	0.06
Log establishments $_{j,-g,t}$	0.61	1.57	1.18	2.24
Year	1994	10	1992	9

The unit of analysis is the establishment (i) –year (t). Firms are indexed by j, while markets (MSAs) are indexed by g. LFP is labor factor productivity.

Table 2: Summary statistics for the entrant datasets

	Hotel sample (n=261,906)		Restaurant sample (n=1,459,118)	
	Mean	Stdv	Mean	Stdv
Sales ($000)	811	4,661	668	934
Collocated entrant dummy (COLO$_i$)	0.06	0.23	0.13	0.33
Payroll ($000)	227	1,373	187	307
Age (years)	4.2	6.0	4.0	6.2
Log establishments$_{gt}$	5.39	1.00	7.44	1.18
MSA LFP$_{-i,gt}$	0.00	0.10	-0.01	0.06
Log establishments $_{j,-g,t}$	0.50	1.42	1.11	2.16
Year	1994	10	1992	9

The unit of analysis is the establishment (i) –year (t). Firms are indexed by j, while markets (MSAs) are indexed by g. LFP is labor factor productivity.

Table 3: Baseline collocation results for incumbent (pre-exiting) establishments

Dependent variable = log sales	Hotel sample		Restaurant sample	
	(1)	(2)	(3)	(4)
$ESTABS_{it}$	**0.024***	**0.021***	**0.008***	**0.011***
	(0.005)	**(0.004)**	**(0.001)**	**(0.001)**
Log establishment payroll	0.722*	0.831*	0.690*	0.697*
	(0.003)	(0.003)	(0.001)	(0.001)
Log establishment age		0.006*		0.003*
		(0.001)		(0.001)
Log $establishments_{gt}$		-0.003		-0.008*
		(0.004)		(0.002)
MSA $LFP_{-i,gt}$		0.064*		0.242*
		(0.009)		(0.011)
Log $establishments_{j,-g,t}$		-0.002		0.004*
		(0.002)		(0.001)
Establishment fixed effects	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y
Ownership change dummy	N	Y	N	Y
Firm LFP dummies (quartiles)	N	Y	N	Y
MSA-Year fixed effects	N	Y	N	Y
Constant	Y	Y	Y	Y
N	180,583	180,583	856,699	856,699
Adjusted-R^2	0.91	0.95	0.87	0.88

* Significant at the 5% level, + Significant at the 10% level.

This table shows how an establishment's dollar-denominated productivity changes with the intensity of within-firm collocation in a particular market g (MSA), where i is an establishment, j is a firm, and t is a year. $ESTABS_{it}$= log $establishments_{it}$= log $establishments_{jgt}$. LFP is labor factor productivity. Standard errors are clustered at the establishment level.

Table 4: Collocation results for new establishments

Dependent variable = log sales	Hotel sample		Restaurant sample	
	(1)	(2)	(3)	(4)
Collocated entrant dummy ($COLO_i$)	**0.041***	**0.024***	**0.038***	**0.020***
	(0.004)	**(0.006)**	**(0.001)**	**(0.004)**
Log establishment payroll	0.926*	0.878*	0.757*	0.769*
	(0.001)	(0.004)	(0.000)	(0.005)
Log establishment age	-0.005*	-0.001*	-0.005*	0.004*
	(0.001)	(0.002)	(0.000)	(0.001)
Log establishments$_{gt}$	0.005*	0.010*	-0.005*	-0.008*
	(0.001)	(0.002)	(0.000)	(0.001)
MSA LFP$_{-i,gt}$	0.066*	0.057*	0.434*	0.209*
	(0.007)	(0.013)	(0.008)	(0.015)
Log establishments $_{j,-g,t}$	0.014*	0.021*	0.004*	0.005
	(0.001)	(0.009)	(0.000)	(0.006)
Firm fixed effects	N	Y	N	Y
Year fixed effects	Y	Y	Y	Y
Ownership change dummy	Y	Y	Y	Y
Firm LFP dummies (quartiles)	Y	Y	Y	Y
Constant	Y	Y	Y	Y
N	261,906	261,906	1,459,118	1,459,118
Adjusted-R^2	0.95	0.95	0.87	0.88

* Significant at the 5% level, + Significant at the 10% level.

This table shows the relationship between an establishment's collocation status upon entry into a particular market g (MSA) and its dollar-denominated productivity, where i is an establishment, j is a firm, and t is a year. Collocated entrant dummy is equal to one if a new establishment entered a market where its parent firm already owned an existing establishment and is zero otherwise. LFP is labor factor productivity. Standard errors are clustered at the firm level.

Table 5: Stock and flow effects of collocation

Dependent variable = log sales	Hotel sample		Restaurant sample	
	(1)	(2)	(3)	(4)
Positive treatment ($COLO+_{it}$)	**0.033***	**0.033***	**0.028***	**0.028***
	(0.004)	**(0.004)**	**(0.001)**	**(0.001)**
Negative treatment ($COLO-_{it}$)	**-0.016***	**-0.014***	**0.011***	**0.011***
	(0.006)	**(0.006)**	**(0.001)**	**(0.001)**
Log establishment payroll	0.831*	0.829*	0.697*	0.697*
	(0.003)	(0.003)	(0.001)	(0.001)
Log establishment age	0.005*	0.006*	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.01)
Log establishments$_{gt}$	-0.003	-0.011*	-0.007*	-0.004
	(0.004)	(0.004)	(0.002)	(0.002)
MSA LFP$_{-i,gt}$	0.064*	0.081*	0.242*	0.162*
	(0.009)	(0.010)	(0.011)	(0.012)
Log establishments $_{j,-g,t}$	0.001	-0.002	0.004*	0.004*
	(0.002)	(0.002)	(0.001)	(0.001)
Establishment fixed effects	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y
Ownership change dummy	Y	Y	Y	Y
Firm LFP dummies (quartiles)	Y	Y	Y	Y
MSA-Year fixed effects	N	Y	N	Y
Constant	Y	Y	Y	Y
N	180,583	180,583	856,699	856,699
Adjusted-R^2	0.91	0.95	0.87	0.87
Stock effect = COLO+ + COLO-	**0.017***	**0.018***	**0.039***	**0.039***
	(0.008)	**(0.008)**	**(0.002)**	**(0.002)**

* Significant at the 5% level, + Significant at the 10% level.

This table shows how an establishment's dollar-denominated productivity changes with positive and negative treatment in a particular market g (MSA), where i is an establishment, j is a firm, and t is a year. Positive treatment is equal to one in the year when a lone establishment becomes collocated with another establishment from the same firm, and is zero otherwise. Negative treatment is equal to one in the year when a previously collocated establishment becomes the lone establishment owned by a firm in a particular market. LFP is labor factor productivity. Standard errors are clustered at the establishment level.

Table 6: Matched sample analysis

Dependent variable = log sales	Hotel sample (1)	Restaurant sample (2)
$ESTABS_{it}$	**0.021***	**0.019***
	(0.004)	**(0.006)**
Log establishment payroll	0.940*	0.852*
	(0.002)	(0.016)
Log establishment age	0.034*	0.007*
	(0.006)	(0.002)
Log $establishments_{gt}$	-0.002	-0.013*
	(0.021)	(0.006)
MSA $LFP_{-i,gt}$	0.144*	0.233*
	(0.043)	(0.044)
Log $establishments_{j,-g,t}$	0.002	0.004*
	(0.004)	(0.002)
Establishment fixed effects	Y	Y
Year fixed effects	Y	Y
Ownership change dummy	Y	Y
Firm LFP dummies (quartiles)	Y	Y
Constant	Y	Y
N	11,873	43,128
Adjusted-R^2	0.95	0.90

* Significant at the 5% level, + Significant at the 10% level.

This table shows a matched sample analysis of how an establishment's dollar-denominated productivity changes with the intensity of within-firm collocation in a particular market g (MSA), where i is an establishment, j is a firm, and t is a year. The matched samples were created using propensity score matching (hotel sample), or Coarsened Exact Matching (restaurant sample), where observations that were positively and negatively treated were matched separately to control group observations 1:1 and then pooled for each sample in the "second stage" analysis in this table. $ESTABS_{it}$=log $establishments_{it}$=log $establishments_{jgt}$. LFP is labor factor productivity. Standard errors are clustered at the establishment level.