NBER WORKING PAPER SERIES

MANAGERS AND PRODUCTIVITY IN RETAIL

Robert D. Metcalfe
Alexandre B. Sollaci
Chad Syverson

Working Paper 31192
http://www.nber.org/papers/w31192

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2023

We thank participants at the Empirical Management Conference, NBER Summer Institute Personnel Economics, European Economic Association Meetings, Duke, MIT, Princeton, as well as Tim Armstrong, Vittorio Bassi, Stephane Bonhomme, Leah Boustan, Thomas Chaney, Oren Danieli, Michela Giorcelli, Bob Hahn, Nathan Hendren, Jonas Hjort, Mitch Hoffman, Matt Kahn, Rem Koning, Simon Quach, Raffaella Sadun, and John Van Reenen for excellent comments. We also appreciate the willingness of the two retailers to share their data for research. We received no remuneration from them for this work. The views expressed in this paper are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, IMF management, or the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w31192

Managers and Productivity in Retail
Robert D. Metcalfe, Alexandre B. Sollaci, and Chad Syverson
NBER Working Paper No. 31192
April 2023
JEL No. D20,L2,M5

ABSTRACT

Across many sectors, research has established that management explains a notable portion of productivity differences across organizations. A remaining question, however, is whether it is managers themselves or firm-wide management practices that matter. We shed light on this question by analyzing store-level data from two multibillion-dollar retail companies. In this setting, managers move between stores but management practices are set by firm policy and largely fixed, allowing us to hone in on managers' personal roles in determining store performance. We find: (i) managers affect and explain a large share of the variance of store-level productivity; (ii) negative assortative matching between managers and stores, which may reflect both firms' decisions and a selection-driven bias that we characterize and argue might apply in other settings using movers designs; (iii) managers who move do so on average from less productive to more productive stores; (iv) female managers are less likely to move stores than male managers; (v) manager quality is generally hard to explain with the observables in our data, but is correlated with the ratio of full-time to part-time workers; (vi) managers who obtain high labor productivity also tend to obtain high energy productivity, revealing some breadth in managers' skills applicability; (vii) high-performing managers in stable growth times are also high-performing during turbulent times; and (viii) exogenous productivity shocks improve the quality of initially low quality managers, suggesting managers can learn. We explain implications of these findings for productivity research.

Robert D. Metcalfe
Department of Economics
University of Southern California
Los Angeles, CA 90007
and NBER
robert.metcalfe@usc.edu

Alexandre B. Sollaci
International Monetary Fund
700 19th St NW
Washington, DC 20431
abalduinosollaci@imf.org

Chad Syverson
University of Chicago
Booth School of Business
5807 S. Woodlawn Ave.
Chicago, IL 60637
and NBER
chad.syverson@chicagobooth.edu

1 Introduction

A growing literature has established that management explains a notable portion of business performance. This work has spanned settings that vary both by industry/sector and the development level of the broader economy the business operates within. Examples of relevant studies include, among others, papers by Ichniowski, Shaw and Prennushi (1997); Ichniowski and Shaw (1999); Bertrand and Schoar (2003); Ichniowski and Shaw (2003); Bloom and Van Reenen (2007); Bloom, Sadun and Van Reenen (2012); Bloom et al. (2013, 2014); Kaplan and Sorensen (2017); Bruhn, Karlan and Schoar (2018); Bloom et al. (2019); Giorcelli (2019); Janke, Propper and Sadun (2019); Bandiera et al. (2020); Gosnell, List and Metcalfe (2020); Sandvik et al. (2020); Bianchi and Giorcelli (2022).

Management's effects on performance can operate through two channels that are not mutually exclusive: (1) the manager herself and (2) firm-level management practices (Syverson, 2011). This is an important distinction. In principle, a firm's mandated management practices can be part of the intangible capital of an organization, and as such are transferable over time and space. Individual manager's performance effects, on the other hand, are inherently tied to the manager. This distinction has implications regarding firms' optimal strategies for improving performance and productivity, the distribution of the rents yielded by superior management, and how government interventions to improve management might increase national productivity levels.

Much of the literature has focused on the managerial practices channel. This is especially true of the experimental literature, as practices are more amenable to manipulation than are managers' personalities, preferences, and innate abilities. Some work has looked at managers' personal effects (such as innate ability and personality styles), though this has focused mostly on corporate CEOs (Barberis et al., 1996; Bertrand and Schoar, 2003; Fee, Hadlock and Pierce, 2013; Bandiera et al., 2015; Bender et al., 2018; Bandiera et al., 2020; Huber, Lindenthal and Waldinger, 2021; Acemoglu, He and le Maire, 2022; Rubens, 2022; Baltrunaite, Bovini and Mocetti, 2023). One potential confound in the corporate CEO effects literature is that it relies on CEOs moving across firms for identification, which opens the possibility that firm-level unobservables may shape the selection of certain managers or their influence on performance.

In this paper, we shed additional light on individual managers' influence on business performance. Unlike the CEO effects literature, our setting allows us to control for firm-level influences on performance, such as governance, corporate culture, pricing, branding, product assortment, supply chain structure, and firm policies regarding management practices. This allows us to focus more on the roles of particular people (managers) rather than practices. Managers in our setting have multiple margins to affect store-level productivity, such as motivating employees; choosing which employees to hire, fire, and promote; enforcing and managing employee schedules and store technology; providing a pleasant physical environment; reducing stock-outs; and implementing company-level changes in price and product assortment.[1] Of course, this list makes clear we do not eliminate all variations in practices across managers, but given our within-firm design we can go considerably farther than most prior studies.

Our data includes information on the performance and management of thousands of stores owned by two multi-billion dollar retail companies, one in the US and one in the UK.[2] Our data are rich enough to not just measure how much individual managers matter, but also to allow us to characterize the mechanisms through which their influence flows into measured performance. This includes, for example, managers' effects on labor efficiency and energy efficiency. We do all of this through operations data rather than surveys, which reduces a host of measurement biases and validity issues.

Our estimation approach allows for heterogeneity not just in managers but in the stores themselves, separate from any manager effects. This allows us to capture how managers might influence stores of different types. It also lets us see whether and to what extent the firm's allocation of managers across stores harnesses any complementarities between store and manager types. In this way, our setting and approach let us zoom out beyond measuring manager heterogeneity and explore more broadly the optimality of manager allocation.

Estimating individual manager and store effects on productivity poses methodological challenges. First, not all manager and store fixed effects can be separately estimated (Abowd, Kramarz and Margolis, 1999). Identification of both effects relies on managers moving between stores. A benefit of working with two large companies is that we have a number of large connected sets through which we can measure many manager fixed effects. Across the two companies, we can estimate 3,584 manager fixed effects and 2,773 store fixed effects. Second, even when those effects can be separately identified, there is

[1]There is some previous research suggesting that management discretion in how some of these are implemented could impact on productivity (Hoffman, Kahn and Li, 2018).

[2]The managers in our data are quite representative of the 8.9 million managers in the United States (US Census Bureau) and the 3.5 million managers in the United Kingdom (Office of National Statistics) with respect to salary, education, and the scope of management responsibilities.

seldom a large number of observations for each individual. This means the estimated fixed effects contain measurement error. As shown by Andrews et al. (2008), this introduces bias into the estimated moments of the fixed effects distribution. To address these biases, we implement the correction procedures described in Andrews et al. (2008) and Gaure (2014). We also apply an Empirical Bayes adjustment where, given prior distributions for measurement error and true manager/store effects, we construct a posterior update for those distributions using the estimated data. The mean of this posterior is the adjusted fixed effect (Chandra et al., 2016).

These adjustments allow us to learn about the joint distribution of manager and store fixed effects. Its shape and range in particular tell us much about the extent to which managers influence stores' productivity. This is another element that distinguishes our approach from previous studies, which have been more concerned with particular cases of measurement error bias, such as attenuation when including estimated fixed effects as regressors (e.g., (Bertrand and Schoar, 2003)). Moreover, another advantage of our inside-the-firm measurement is that we can abstract away from across-firm heterogeneity, which Sorkin (2018) and Di Addario et al. (2022) have shown to be large.

We obtain eight major findings that are remarkably consistent across the two independent companies. First, individual managers matter for productivity. While store fixed effects are somewhat more diffuse and explain a greater fraction of productivity variation, managers still explain an important share. Overall, managers explain between 25 and 35% of the variance of store-level productivity, which is about 50-70% of the explanatory power of store fixed effects. In the four largest connected sets across both companies, moving a manager from the 10^{th} percentile to the 90^{th} percentile increases overall productivity by between 22% and 82%. On average, this implies an effect on output equivalent to adding a fifth employee to a team of four.

Second, manager and store effects tend to be negatively correlated in our data. While in contrast to our expectations, we posit three reasons for this negative assortative matching (NAM). First, taken at face value, this NAM result could mean that there are no substantial complementarities between managers and stores (Becker, 1973), and there is some long-term advantage for the companies in placing high-performing managers in low-performing stores. Second, there might be a benefit of positive assortative matching, but firms are not aware of it and do not take advantage of it in their manager allocation decisions. Third, we show that productivity-based selection on manager-store pairings may negatively bias the estimated covariance, perhaps clouding true complementarities.

We show that if the likelihood of a manager-store breaking up (resulting in the manager moving) decreases as a positive function of manager and store types, this induces a negative correlation between manager and store fixed effects among movers. Thus even if the true correlation across the entire population of manager-store pairs is zero or positive, it may be estimated as negative. We are unaware of prior research suggesting this selection-based bias toward finding NAM, but we believe that it has broader implications for empirical tests of complementarities when equilibrium pairings of inputs are subject to performance thresholds. This potentially includes many of the empirical settings in the two-way fixed effects literature using AKM-type (Abowd, Kramarz and Margolis (1999)) estimators.

Third, there are meaningful changes at the manager and store level during manager changeovers at stores. In one company in our dataset, old managers' earnings fall in the months prior to their replacement. At the other company, stores' productivity and sales fall before manager departure, but this reverses once the new manager arrives. Replacement store managers tend to be newer to the firm; they average almost 17 months shorter tenure in one company, and 6.5 months less in the other. Despite those differences, incoming managers are not systematically better or worse than outgoing managers by our metric. The distribution of changes in managerial quality (the estimated fixed effect) is centered around zero, and its average is not statistically distinguishable from zero. In contrast, managers who move between different stores in the same retailer tend to move to better (higher fixed effect) stores. In fact, managers that start in a store in the bottom of the quality distribution are more likely to leave that store than managers who start in more productive stores. In addition, when those managers leave, they most likely go to a higher-type store. Those movements are observed in both companies, though are more pronounced in the American firm.

Fourth, women are less likely to be movers within the company. This is not due to tenure, features of their initial store (size, revenue, number of employees, format), or the quality of the store-manager match. We find a substantive difference in propensities to move between female managers in large cities with multiple stores and those in smaller, single-store cities or towns. In large cities, women managers are much more likely to move and progress through the ranks. This result is consistent with the literature on how women's family responsibilities often constrain them to a single geographical area, limiting their career paths (Le Barbanchon, Rathelot and Roulet, 2021).

Fifth, we characterize which observable features in our data are correlated with manager and store quality. We find a statistically significant relationship between manager quality and the ratio of full-time to part-time workers in their stores. However, the sign of the effect is different between the two companies (positive in the US and negative in the UK). Most observables in our data do not have a statistically significant association with manager quality, including the manager's tenure, gender, distance to the nearest competing store, and even wages. Thus a large part of the variation in manager quality remains unexplained, suggesting that less measurable characteristics of the manager (e.g., leadership style and charisma) could play an important role.[3] As to store quality, higher quality stores tend to be in locations with higher average incomes, pay higher wages (which could explain the average increase in store quality when managers switch between stores in the same company), and as with the manager fixed effects are correlated with the ratio of full-time to part-time workers, albeit again with different signs between companies. Better stores are also slightly more energy efficient. Distance to competitors, store size (floor area), and manager gender have no effect on store quality.

Sixth, manager quality and energy efficiency are positively correlated. More productive managers use less energy; a 1 p.p. increase in manager quality is associated with a 0.11 p.p. decrease in energy use, holding fixed store size, total sales, energy source, location, temperature, time of year, energy-efficient capital, and electricity prices. This suggests that good managers are like TFP, in that they raise the productivity of multiple inputs. This result has some similarities with the literature suggesting that energy and capital are complements in the production function (Berndt and Wood, 1979; Pindyck and Rotemberg, 1983; Atkeson and Kehoe, 1999). What we show is that managers might be an important mediator for that complementarity.[4] As in several other studies in the weather-energy use literature (Deschênes and Greenstone, 2011; Auffhammer and Mansur, 2014), we find temperature has a U-shaped impact on energy consumption. More productive managers can attenuate this U-shape, resulting in smaller increases in energy consumption during the colder and warmer months (beyond their baseline impact on energy use).

[3]Consistent with the literature on agglomeration effects (Duranton and Puga, 2004; Combes et al., 2012; Gennaioli et al., 2013; Behrens, Duranton and Robert-Nicoud, 2014), we find a positive relationship between population density and store-level productivity. Across both companies, we find that such effects are coming through the manager fixed effects and not the store fixed effects. This suggests a sorting of workers to areas is driving the agglomeration effects within these two companies. We also show novel evidence that agglomeration effects can be observed within firms, not just at the cross section of cities, and that management is the driver of effects.

[4]We have individual store-level data on energy use and prices, different from the literature that leverages state-level price changes or uses use a composite energy index (hiding a lot of important heterogeneity across fuels) (Popp, 2002; Kahn and Mansur, 2013; Aldy and Pizer, 2015; Marin and Vona, 2021). Notable exceptions include Davis et al. (2013); Singer (2022). Moreover, in our case the heterogeneous energy price changes are not necessarily passed through to consumers, because pricing for goods is done centrally (as is the case for many large companies (DellaVigna and Gentzkow, 2019)). Thus we worry less that these changes affect local prices and demand.

High-performing managers are thus good for climate change mitigation.

Seventh, we compare the explanatory power of managers on productivity in normal versus turbulent economic times (i.e., the Covid-19 pandemic). While research has investigated how companies have responded to disruptions (Decker et al., 2018; Aghion et al., 2021) and during the Covid-19 pandemic in particular (Bell, Bloom and Blundell, 2021; Bloom et al., 2020; Andrews, Charlton and Moore, 2021; Bloom, Fletcher and Yeh, 2021; Barry et al., 2022), no previous work has examined managers' behavior and performance in response to such shocks. We find that managers with high fixed effects (estimated using only data during stable pre-Covid-19 periods) also performed better than average during the Covid-19 period.

Eighth, using detailed store-level data on exogenous disruptions before and during Covid-19, we find disruptions have a strong negative impact on productivity, but higher-quality managers mitigate these effects. However, managers of any type exposed to more disruptions pre-Covid were better at limiting Covid-related productivity disruptions. This was especially true for low-quality types, raising the possibility that lower quality managers learn more from experience with previous shocks, perhaps because they have more room to grow in the first place. This is new evidence that suggests manager fixed effects might change over longer periods (echoing the results in Lachowska et al. (2022) with company-level fixed effects changing over time).

There are several closely related papers to our first result, including Lazear, Shaw and Stanton (2015), Janke, Propper and Sadun (2019), Fenizia (2022), and Giardili, Ramdas and Williams (2022). Lazear, Shaw and Stanton (2015) shows that supervisors of computer-based test graders matter, where 10th-to-90th percentile supervisor difference has an equivalent effect on output as adding one worker to a team of 9. Janke, Propper and Sadun (2019) shows hospital managers do not measurably influence health-related productivity metrics. Fenizia (2022) uses a movers design and shows that within the Italian social security administration, manager fixed effects explain 9% of the total variation in claims productivity at the office level. Giardili, Ramdas and Williams (2022) examine car manufacturing plants and managers and show that managers explain about 7% of production differences. Our estimate of the effects of managers on productivity is much larger than these results. This difference could reflect in part our retail setting, where it is plausible that managers of customer-facing stores have more scope to influence realized productivity. Also, these other papers did not use revenue as their output metric, which may capture more channels through which managers affect store performance. Further-more, we have more and larger connected sets of managers and stores across both of our two companies than these previous papers.

Our research speaks to existing work on the margins that managers can improve the productivity of high- and low-performing workers (Mollick, 2012; Lazear, Shaw and Stanton, 2015; Best, Hjort and Szakonyi, 2017; Adhvaryu, Kala and Nyshadham, 2019; Frederiksen, Kahn and Lange, 2020; Patault and Lenoir, 2020; Hoffman and Tadelis, 2021; Limodio, 2021; Cai and Wang, 2022; Friebel, Heinz and Zubanov, 2022; Minni, 2022). It also has ties to the matching literature in labor economics (Eeckhout and Kircher, 2018) and more specifically to the negative assortative matching result in Card et al. (2018), and Adhvaryu et al. (2020) who examine how managers choose workers to increase productivity. It is also related to choosing social movement (union) leaders on preferences for wage changes (Boudreau et al., 2021).

We extend the two-way fixed effects sorting models of labor across firms (Abowd, Kramarz and Margolis, 1999; Card, Heining and Kline, 2013; Lopes de Melo, 2018) by examining how the manager impacts on productivity and not just whether their wages change in the new firm/location. This is important because productivity is of inherent interest, wages do not always map well onto productivity, and there are a range of identification and measurement issues using wage data (Eeckhout and Kircher, 2011; Hagedorn, Law and Manovskii, 2017). Such measurement and identification issues do not exist when revenues and costs are observed within each unit that the manager manages. On a broader methodological front, as we note above, we show that one of the potential reasons for the observed negative assortative matching between managers and stores results from a selection-based bias in mover identification strategies (i.e., when the likelihood of a manager-store breaking up decreases in a function of the manager and store types). The reliance on two-way fixed effects estimation on movers for identification means this might bias estimates toward negative assortative matching. While we elucidate this selection-based bias in our setting, it may apply apply to all two-way fixed effect estimation settings with mover designs (worker-company, person-neighborhood, etc.) where the stability of pairs is an increasing function of both types. We believe we are the first to demonstrate this issue and is separate from the other biases in the two-way fixed effect literature, such as limited mobility bias (Andrews et al., 2008).

In the remainder of this paper, Section 2 highlights the data from the two companies, and Section 3 then estimates the manager and store fixed effects. Section 4 goes into depth on explaining the causes and effects of managerial changes within the firms. Section 5 calculates what explains manager and store productivity. Section 6 then uses the Covid-19 shocks to understand how manager effects mediate responses to the pandemic. Section 7 concludes.

2 Data

Our dataset comprises monthly operations data at the store level for two large retail companies. Data for Company A, an American company, is available between April 2018 and December 2020 (we divide this sample into pre- and post-Covid-19 periods with the threshold being February 2020; Covid-19 seriously disrupted the company's regular operations). Data for Company B, a British company, is available between April 2014 and May 2017.

In both cases, the data includes thousands of stores located across practically the whole territory of their respective countries. For each store, we observe monthly sales, full- and part-time employment, store size (floor area), location (city and state/region), some other factor inputs, and a set of additional store characteristics. Importantly, we also observe the identity of the manager in each store, which allows us to track managers as they move across establishments over time.[5]

Stores change managers relatively frequently, offering us reasonable leverage to identify manager effects. In both companies, more than half of stores change their manager at least once during our sample. Of those, about 60% change managers only once, 30% change managers twice, and the remainder more frequently than that. Similarly, 14 or 23% of managers (depending on the company) work in two different stores, and a small fraction move through three or more. Between 20 and 30% of each company's manager changes are internal, that is, involve moving a manager between different stores of the same company.

2.1 Movers and Non-movers

We do not observe why managers leave their stores (this is explored further in section 4) and therefore cannot rule out the possibility that manager departures are related to store characteristics. Nevertheless, we see a significant overlap between the distributions

[5]We drop the handful of stores in each company that have multiple managers within a given month.

of outcomes across mover and non-mover stores (see Table B.1). In addition, Figure A.1 shows that average outcomes within these two groups of stores follow very similar paths over time, despite differences in levels (Appendix E contains a more detailed discussion on pre-trends).

One clear difference across mover and non-mover managers is gender.[6] Specifically, women are about 7% less likely to move to another store than are men. This difference cannot be explained by manager tenure nor the characteristics of the initial store (where the manager moves from) such as number of employees, revenue, size, or format (flagship, local store, etc). This gap becomes weaker and loses statistical significance when we include only managers that do not leave the company (i.e., are still present in the final period of the dataset), but it increases again when we further constrain the data to include only managers whose first store was in a city with at least 10 stores of the same company (Table B.2).[7] When they do move, the average distance moved by men and women is very close (27 vs 25 miles).

A possible explanation for this finding is that women are less likely to move due to family reasons. If promotions within the company are associated with moving to different stores, female managers might instead find it more attractive to move to a different company. This would explain why the difference disappears when we condition on managers that do not leave, but strengthens when only large cities are included, where managers could more easily change stores without changing residences. While inherently interesting, this finding is unlikely to interact with the results in the following sections, as gender and productivity do not seem to be related.

3 Manager and Store Effects

In this section we detail how we decompose store productivity into manager and store fixed effects. We discuss the identification of those effects and the steps we take to mitigate estimation error.

[6] Gender is inferred from the manager's title (Miss, Mr, Mrs, etc.). In the small number of cases where the title is insufficient to infer gender, we set it as missing.

[7] This result resonates with Dauth et al. (2022), where large cities facilitate better matching.

3.1 Measuring Productivity

We measure store-level productivity in each period as revenue (sales) per full-time equivalent employee:

$$prod_{s,t} = \frac{sales_{s,t}}{employment_{s,t}}.$$

There are two main benefits from adopting this definition. First, revenue per employee can be directly constructed in the data and is intuitively related to less observable primitives driving stores' performance levels. Second, it is a widely used metric and likely to reflect the performance indicators that managers and companies target. In this sense, our productivity measure may well be more aligned with managers' incentives than underlying demand or supply primitives.

Like all productivity measures, this metric is a residual. Any variation in revenues unaccounted for by labor inputs or other controls will be in our productivity measure. This might include variations in stores' physical capital, locations, or nearby competitors. In section C.1 of the appendix, we compare sales per employee with other potential measures of productivity, such as sales per store area (which proxies for physical capital), and revenue-based TFP (which controls for all observable inputs). We find all measures are highly correlated and appear to capture the same underlying store characteristics.

3.2 Decomposing Productivity

We decompose each store's productivity into three components: a store-specific component, a manager component, and a time (month) component. Formally, we estimate the fixed effect model

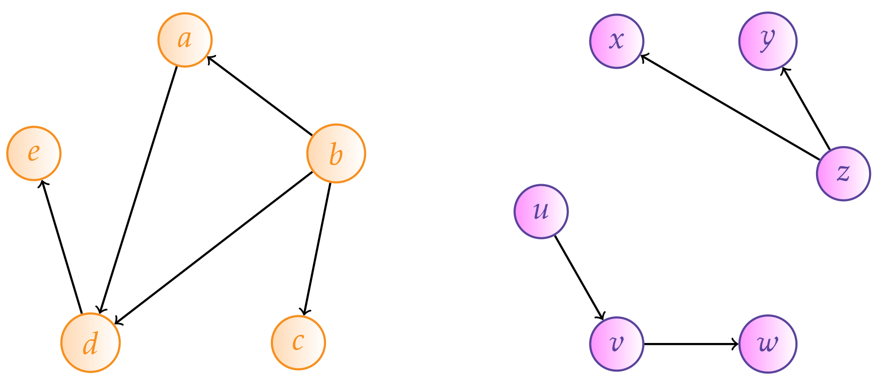
$$\log(prod_{s,t}) = \mu_s + \mu_{m(s,t)} + \mu_t + \varepsilon_{s,t} \tag{1}$$

where μ_s is the store fixed effect, $\mu_{m(s,t)}$ the fixed effect of the manager associated with store s and time t, μ_t a time fixed effect, and $\varepsilon_{s,t}$ the residual. The store fixed effect in this model captures any time-invariant influences on store productivity, be it physical size, location, type of establishment, local competition, or other possibilities. The manager fixed effect captures the manager average effect on the productivity of the stores where they worked. Throughout the paper, we interchangeably refer to the manager's fixed effect as their productivity, ability, or quality. The time fixed effect absorbs any company-wide productivity shocks.

Note that the log-linear regression specification (1) imposes assumptions on the complementarity between the manager and store components of productivity. Specifically, it implies that the level of productivity is supermodular in the store and manager fixed effects (assuming their elasticities are positive). Section C.2 of the appendix digs into this issue, showing that more general specifications produce results very similar to our baseline.

Estimation of this fixed effects structure brings a series of challenges. First, not all store and manager fixed effects can be separately identified. This point can be better understood if we think of μ_s and $\mu_{m(s,t)}$ as the coefficients on a complete set of dummies for stores and managers. If a store has the same manager for the duration of our sample, there is no way to separately identify each fixed effect, as the manager and store dummies are perfectly collinear. The approach therefore relies on managers that move across stores for identification.

Figure 1: Examples of Store Networks



(a) One connected set (b) Two connected sets

The network structure of stores also matters for identification. Figure 1 shows two examples of store networks. Each circle represents a store, and a link between them means the stores have shared a manager (for example, a manager that worked at store a moved to store d). Network (a) (containing stores a through e) has a single *connected set*, meaning that all stores in the network have, directly or indirectly, shared at least one manager. Because of that, we can estimate the fixed effects of managers and stores relative to a common baseline (the excluded dummy) and therefore meaningfully compare the quality of managers and stores within that network.

In contrast, network (b) (stores u through z) has two separate connected sets, $\{u,v,w\}$

and $\{x,y,z\}$. Consider, for example, a firm that has establishments in both the East Coast and the West Coast of the US. Suppose managers are happy to move to different stores along the same coast but unwilling to move across the country. This would create the pattern in network (b), where stores in each coast would be connected to each other, but no managers connect stores across coasts.[8]

Because there are two separate connected sets in network (b), the estimation of fixed effects requires the normalization of two separate baselines—one omitted dummy per connected set. And because the omitted dummy is arbitrary, the fixed effect of store x, for example, cannot be compared to the fixed effect of store w, because they are measured relative to different, unobserved baselines. This means that we can only compare magnitudes of fixed effects inside the same connected set (see also Abowd, Kramarz and Margolis, 1999; Fenizia, 2022).

Our data contains about 890 connected sets in Company A and just over 1,200 connected sets for Company B. The vast majority of these are trivial, including just a single store and manager (i.e., managers that do not move). The four largest connected sets in Company A have between 13 and 18 managers and 7 to 10 stores. The sets are much bigger in Company B. The largest set has 202 managers and 130 stores. Set sizes to drop quickly, however, with the next 3 largest connected sets respectively containing 100, 66, and 55 managers and 67, 44, and 32 stores. In many of our results below, we focus only on the largest connected sets of each company as those allow for more precise estimates. Figure A.4 in the appendix summarizes the structure of the connected sets in our data.

A second estimation challenge in model (1) is that the number of parameters can be very large. This poses a computational problem, as regular OLS estimation methods become slow and memory-intensive. Fortunately, a number of algorithms have been developed for this purpose. We use the `lfe` package in R (Gaure, 2013).[9]

The third major challenge in our framework involves estimation error, as our fixed effects are typically estimated with relatively few observations. Estimation error causes issues in the analysis of the fixed effects. For example, even if measurement error has mean zero and is independent from the true parameters, it can still introduce bias into the variance of our estimates, as well as to the covariance between store and manager fixed effects, sometimes referred to as limited mobility bias (Andrews et al., 2008). We

[8] This is like what we observe in the data for Company A: the overwhelming majority of connected sets includes only stores in adjoining states.

[9] For consistency throughout the estimation process, we estimate fixed effects using the fixed-point iteration algorithm, implemented in Stata by Correia (2017).

address this below.

Lastly, there is also the possibility that the estimates of the fixed effects μ_s and μ_m are endogenous. Specifically, if managers are more likely to move to (or from) a given store s after that store experiences a positive (negative) shock in productivity, the residual $\varepsilon_{s,t}$ could be correlated with the manager fixed effect. Note, however, that this requires the likelihood of leaving/joining a store to be correlated with the manager quality itself. If, for example, all managers are equally likely to move away from a store after a negative productivity shock, there is no correlation between $\varepsilon_{s,t}$ and $\mu_m(s,t)$. While we cannot directly address those concerns in this paper, we have found that the decision to move/stay in a store is seldom made by manager alone. Staffing rules and company needs – all uncorrelated with productivity shocks – also play an important, if not the most important, role in those decisions. In fact, as we will see below, good managers tend to be matched to bad stores, which goes against the endogeneity argument sketched above.

3.2.1 Variance Decomposition

The relationship between the store- and manager-specific components of productivity is a major revelation of the productivity decomposition. We can study this relationship by looking at the variance of both sides of (1):

$$\text{Var}(\log(prod_{s,t}) - \mu_t) = \text{Var}(\mu_s) + \text{Var}(\mu_{m(s,t)}) + 2 \times \text{Cov}(\mu_s, \mu_{m(s,t)}) + \text{Var}(\varepsilon_{s,t}). \tag{2}$$

The left-hand side of this equation adjusts the variance of logged productivity for time effects, while the right-hand side splits this variance into the variances of the store and manager fixed effects, their covariance, and the variance of the residual (assumed to be exogenous). Two things are of primary interest in this equation: (1) how much of the variance of productivity can be attributed to each component, and (2) the sign of the covariance between the manager and store fixed effects.

Recall from our discussion of estimation issues that this decomposition only has meaning inside connected sets, as manager and store effects are normalized to different levels across sets. In addition, if our estimates of μ_s and μ_m contain measurement error, their variances will be upwardly biased. Further, although this might not be immediately clear, their covariance will be negatively biased (Andrews et al., 2008).[10] We correct for this bias using the method described in Gaure (2014).

Table 1: Decomposition of the Variance of Productivity

CS Rank	Obs		Company A	
		$\frac{\text{Var}(\mu_s)}{\text{Var}(\log(prod_{s,t})-\mu_t)}$	$\frac{\text{Var}(\mu_m)}{\text{Var}(\log(prod_{s,t})-\mu_t)}$	$\frac{\text{Cov}(\mu_s,\mu_m)}{\text{Var}(\log(prod_{s,t})-\mu_t)}$
1	215	0.72	0.05	-0.07
2	191	1.80	1.01	-1.14
3	171	0.21	0.12	0.15
4	142	0.48	0.07	-0.02
CS Rank	**Obs**		**Company B**	
		$\frac{\text{Var}(\mu_s)}{\text{Var}(\log(prod_{s,t})-\mu_t)}$	$\frac{\text{Var}(\mu_m)}{\text{Var}(\log(prod_{s,t})-\mu_t)}$	$\frac{\text{Cov}(\mu_s,\mu_m)}{\text{Var}(\log(prod_{s,t})-\mu_t)}$
1	3,710	1.64	1.49	-1.22
2	1,979	0.73	0.22	-0.14
3	1,179	1.03	0.77	-0.62
4	1,008	0.39	0.30	-0.03

Columns 3-5 in this table show the ratio of the variance in the store FE (or manager FE, or their covariance) and the variance in total productivity according to equation (2). Results are computed within the four largest connected sets for each company.

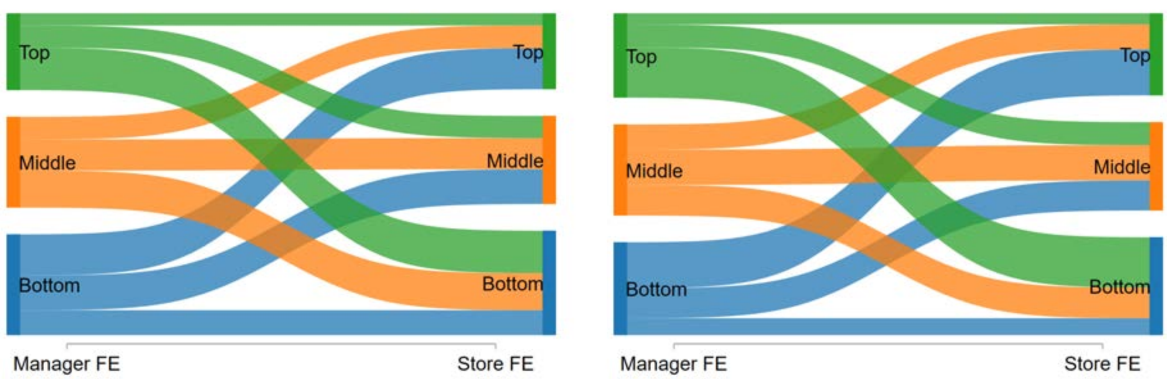
Table 1 shows the ratio of the store fixed effect (FE) variance, manager FE variance, and their covariance to the variance in productivity (after time effects are removed) within each of the four biggest connected sets (CS). There are two key takeaways from this table. In every case the store FE variance is larger than the manager FE variance. Thus stable store-level factors (including local demographics, competition from nearby retailers, etc.) explain a greater share of the observed productivity variation than do managers. At the same time, however, the average amount of the variance manager fixed effects does explain is sizable. Across the connected sets on the table, the manager fixed effect variance is on average 58% of the store fixed effect variance and 50% of the total variance, suggesting an important role for managers to influence outcomes.

The second takeaway is that the covariance between manager and store effects is almost always negative. This can also be seen in the “raw" data by plotting the relative frequency of matches by manager and store quality. As shown in Figure 2, very few top

[10]Intuitively, the bias in the covariance comes from the fact that when the measurement error in the manager fixed effect is positive, the measurement error in the store fixed effect will tend to be negative, and vice versa, as the two components sum to the same total. This conclusion is specific to our setting, however, where there are no other covariates in the model.

managers work in top stores (and equivalently for bottom managers at bottom stores). Instead, the most common match is between a top manager and bottom store, or a bottom manager and top store. Given that our structure implies persistent store and manager attributes are complements (log productivity is additive in the fixed effects, implying a multiplicative effect in levels and hence a positive cross-derivative of their marginal products), this is somewhat puzzling, as matching models would predict assortative matching between heterogeneous complementary inputs if productivity maximization were the objective (e.g., Becker, 1973).[11] We estimate that Company A and Company B could increase their sales revenue (and therefore productivity) by 2.2 and 6.1 percent (USD 2.2 million and GBP 13.2 million) each month, respectively, by reallocating their managers across stores in accordance with positive assortative matching (see Appendix D for details). The negative covariance we find exists in both our sample companies, and is encountered when Fenizia (2022) investigates the matches between heterogeneous-quality Italian government offices and managers. This suggests other mechanisms might be responsible for what we observe. We discuss some possible explanations here.

Figure 2: Match Between Managers and Stores



(a) Company A (b) Company B

This figure shows the matches between managers and stores, split by manager and store quality. In each case, managers and stores are divided into terciles based on their position in the distribution of fixed effects (EB-adjusted estimates of the fixed effects are used for this exercise; see section 3.3). The size of each link in the chart is proportional to the share of matches it represents.

Note that while managers and stores are split into equally sized groups, the chart represents the number of *matches* between groups. Thus, if managers in one group are more less to move (and hence are part of fewer matches), that group will represent a lower share of the total number of matches. This can be seen, for example, in the initial share of bottom and top managers in company A, where top managers represent a smaller share of the total number of matches in the chart.

[11]We estimated a separate specification for log productivity that included the interaction of the fixed effects and found that it has a positive coefficient, providing further evidence of complementarities.

Lack of Awareness. One possible explanation for the negative correlation between the qualities of stores and managers is that companies are simply unaware that they are not assortatively matched, as neither type is directly observable. For example, a related paper by Cowgill et al. (2021) shows that the firm's (or CEO's) preferred match is negatively assortative, but a self-organized match is positively assortative. Given that central HR and regional managers have a strong role in determining the match, this might explain our result. This unawareness might also reflect the findings of previous research suggesting that companies do not always promote the best people to management (Lazear, 2004; Benson, Li and Shue, 2019). Either story suggests companies would benefit from measuring and learning more about managerial quality.

Learning & Reducing Productivity Variation. Another possibility is that our model is incomplete or misspecified. If the *true* production function does not exhibit complementarity between managers and stores, positive assortative matching is not the optimal configuration of managers and stores. For example, the companies might assign better managers to worse stores to reduce productivity variation across locations. This would be desirable if risk-sharing priorities mattered (e.g., Adhvaryu et al., 2020), if they wanted to achieve a minimum standard across all locations (e.g., Fenizia, 2022) for branding purposes, or if they viewed store closures as particularly costly.

Selection-Based NAM. Finally, the negative assortative matching (NAM) we observe across managers and stores could reflect a sample-selection effect arising from way store and manager types affect the likelihood that a store-manager match remains stable. Let us explain.

Suppose that a firm is able to measure the (joint) productivity of a store, but that separating this measurement into a manager and store-specific contribution is costly. A plausible and simple rule that a company might use to select "good" matches is $\mu_m + \mu_s \geq k$, for some arbitrary constant k (i.e., if it's not broken, don't fix it).[12] All other manager/store pairs are unstable matches whose managers will eventually move or be moved. It turns out that this rule alone would create a downward bias in the estimated

[12] Indeed, this selection rule implies two correlations that are consistent with elements of our data. First, it implies that there is negative relationship between the probability that a manager moves and the log-productivity of the store (after controlling for time fixed effects). We find this relationship is negative and significant for company A, but statistically indistinguishable from zero in company B. Second, the selection rule implies that the average quality of managers in good stores is increasing over time (since good managers and good stores form a stable match). Figure A.5 shows that this is true in company B, although there is no clear pattern in company A.

correlation between manager and store quality.

To understand this point, let us analyze the implications of this meets-a-threshold rule for match preservation. In stable matches, we have $\mu_m \geq k - \mu_s$. As a result, for a given store quality m_s, the expected value of a manager's quality in a stable match decreases with m_s: $\mathbb{E}[\mu_m|\text{stable match}] = \mathbb{E}[\mu_m|\mu_m > k - m_s]$. That is, better stores will "tolerate" worse managers on average in a stable match. Poor stores, on the other hand, only achieve a stable match if the manager fixed effect is very high. This conditionality drives down the correlation between the quality of managers and stores in a stable match relative to the correlation among the entire manager and store populations. Of course the same negative correlation holds true within unstable matches, where $\mu_s + \mu_m < k$. Because the estimation approach relies exclusively on managers who move (i.e., that were part of an unstable match), this conditionality results in a negatively biased estimate of the correlation between manager and store fixed effects. Note that this restriction can also affect the range of the observed store/manager qualities, which could bias their variance as well.[13]

Figure 3a illustrates this mechanism by simulating an example where managers and stores are initially randomly matched to each other. As seen in the figure, if all matches are observed, the correlation between manager and store productivity is zero. However, note what happens when we split the sample between stable (blue circles, $\mu_m + \mu_s \geq 0$) and unstable matches (red x's, $\mu_m + \mu_s < 0$). Even though manager and store types are uncorrelated overall, there is now NAM between managers and stores *within* the separate groups of stable and unstable pairs. Given that the two-way fixed effect estimation strategy can only separately identify the productivity of managers and stores for pairs that include movers—in other words, the managers and stores identified by the red x's in Figure 3a—this will create a negatively biased estimate of the overall correlation of store and manager types.

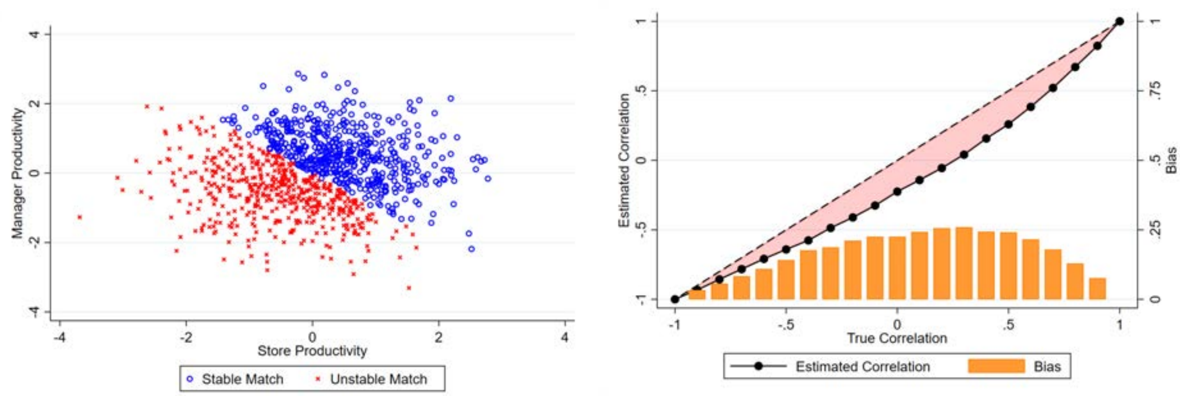
The same result obtains if we reshuffle managers and stores in initially unstable matches and form a second round of matches, and repeat this process indefinitely or until all

[13] A competing explanation is that stores have varying degrees of market power, and because fixed effects are estimated based on revenue, stores with higher market power would appear to be more productive. In that sense, the negative correlation between manager and store fixed effects could instead be a consequence of differences in market power (e.g., Bao, De Loecker and Eeckhout, 2022). This mechanism is unlikely to drive our findings for two reasons. First, most pricing decisions in our sample companies are made at the corporate level, so revenue differences across stores mostly represent differences in quantities rather than prices. Second, we don't find a strong relationship between the measured store fixed effect and the distance to its closest competitors (Figure 9b). If competition were a defining factor for store productivity, we would expect that stores that are closer to competitors would do worse in revenue terms than stores facing no local competition.

matches are stable. Thinking back to the covariances estimated in table 1, we find that they resemble weighted average of the covariances between the manager and store pairs in unstable matches during the first round, and the covariances observed across all managers and stores in each subsequent matching round, where weights are given by the duration of each round.[14] Because the covariance in the first round (which only includes unstable matches) as well as the covariance between stores and managers in stable matches for all subsequent rounds is biased downwards, the overall estimated covariance is also biased downwards.

Figure 3b generalizes the simulation above by allowing for an arbitrary correlation between managers/stores in the initial match and again when reshuffling of managers and stores from unstable matches. The bias is always negative, achieves its maximum value when the correlation is approximately 0.3, and equals zero when the true correlation between manager and store productivity is 1 or -1. In the last case, $\mu_m + \mu_s$ would always lie on a straight line, so imposing $\mu_m + \mu_s \lesseqgtr k$ does not affect the correlation between the fixed effects.

Figure 3: Random Matches of Managers and Stores



(a) Random Matches

(b) Match Correlation and Size of Bias (after two rounds of matches)

Panel (a) shows the distribution of μ_s and μ_s when both are i.i.d. $\mathcal{N}(0,1)$ and $k = 0$. Panel (b) allows for an arbitrary correlation ρ, so that

$$\begin{pmatrix} \mu_m \\ \mu_s \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$$

and there are two rounds of matching. The correlation is estimated by stacking all pairs of managers and stores (across all rounds) and computing the correlation between the two stacked series.

[14]More formally, it is the covariance between a sequence $\{\mu_{m,t}\} = \{\mu_{1,1}, \mu_{1,2}, \ldots, \mu 1, T, \mu 2, 1, \ldots, \mu_{M,T}\}$ of manager fixed effects (m indexes managers and t indexes time) and a similar sequence for matching store fixed effects, where all managers and stores in stable matches during the first round are removed. Because the fixed effects are constant across time, each pair (μ_m, μ_s) is effectively weighted by how long it lasts in the sample.

While we use in the discussion above a deterministic threshold rule for ease of exposition, the logic here applies equally well to cases where match stability is a stochastic function of manager and store types. Moreover, the function need not be linear; any function where stability is increasing in both manager and store types will pose the same selection-based downward biases in variances and covariance.

We note that the selection-based NAM issue we point out here is different from limited mobility bias and, to the best of our knowledge, has not been discussed elsewhere. This, even though it potentially applies in many situations that rely on movers to conduct such two-way fixed effect decompositions. The limited mobility bias described by Andrews et al. (2008) can be understood as an incidental parameter issue, where a large number of firm-specific parameters is identified by a limited number of movers that go through that firm (see also Bonhomme et al., 2020). This bias can be reduced by increasing the number of workers/managers that go through that firm/store. This is not the case in the stable-match-threshold setting we describe above. Increasing the sample size changes nothing about the selection of stable matches between managers and stores. It is also different from endogeneity bias, which arises when managers choose to move to a store because of unobservable characteristics of that store. As shown in Figure 3b, the *observed* correlation between manager and store quality is biased regardless of how managers are assigned to stores.

3.3 Correcting Measurement Error

The correction procedure adopted in the previous section accounts for the bias induced by measurement error on the variance-covariance matrix of the fixed effects, but does not recover the effects themselves. In this section, we focus on methods to recover the true fixed effects by reducing the measurement error of our estimates. We briefly describe the Empirical Bayes adjustment technique used to recover fixed effects from noisy data (see Chandra et al., 2016). Section C.3 of the appendix introduces an alternative method that reduces measurement error by abandoning the estimation of individual fixed effects and focusing on groups of similar managers and stores (Bonhomme, Lamadon and Manresa, 2019).

To implement the Empirical Bayes (EB) adjustment we assume that because of measurement (or estimation) error, the estimated fixed effects are random variables that are normally distributed around their true value. Formally, for each manager m and store s, let $\hat{\mu}_m$ and $\hat{\mu}_s$ denote the estimated effect and μ_m^* and μ_s^* the true effect. Then

$$\hat{\mu}_m|\mu_m^*, \sigma_m \sim \mathcal{N}(\mu_m^*, \sigma_m^2) \text{ and } \hat{\mu}_s|\mu_s^*, \sigma_s \sim \mathcal{N}(\mu_s^*, \sigma_s^2)$$

independently in both cases. Note that independence in this context means the measurement error in a manager's estimated fixed effect is independent only from the measurement error on other *managers'* fixed effects. It does not imply that the measurement error in a manager effect is independent from the error in their associated store effect.

The prior distribution of true manager and store productivity is

$$\mu_m^*|m_m, \tau_m \sim \mathcal{N}(m_m, \tau_m^2) \text{ and } \mu_s^*|m_s, \tau_s \sim \mathcal{N}(m_s, \tau_s^2),$$

independently in either case. We assume the underlying means m_m and m_s are constants, although the method allows for them to be functions of manager and/or store characteristics as well. By conditioning on the estimated fixed effects $\hat{\mu}_m$ and $\hat{\mu}_s$, we can find the posterior distributions for μ_m^* and μ_s^*. The EB-adjusted fixed effects, μ_m^{EB} and μ_s^{EB}, are the means of those posterior distributions.[15] Quantitatively, we find that the EB-adjusted and unadjusted fixed effects are highly correlated. Figure C.5 compares how each of our measures of store and manager quality is related to productivity.

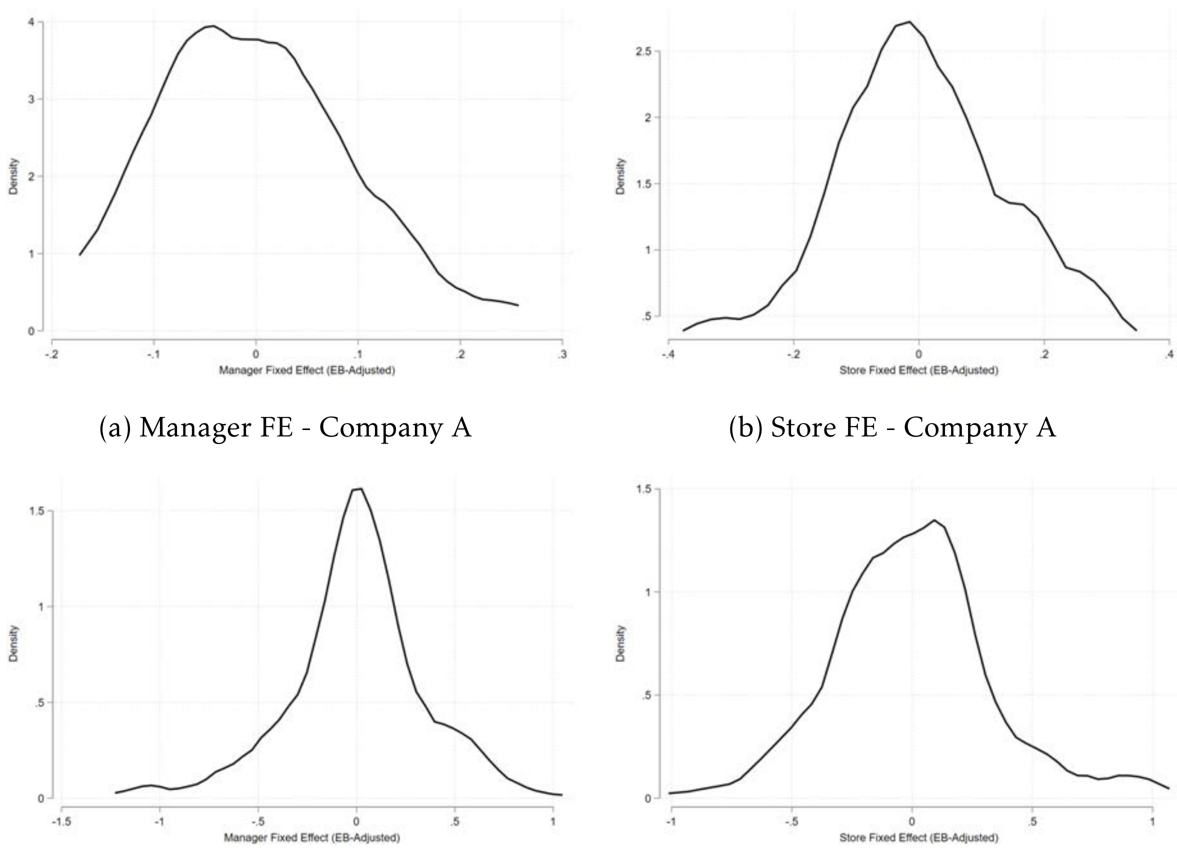
Our results indicate managers can have a significant influence on stores' productivity levels. This can be seen in Figure 4, which plots the distributions of the EB-adjusted store and manager fixed effects inside the largest connected set for each company. In each panel, the fixed effects have been demeaned within their respective connected set, so we focus on the range of the distributions. Note that the distributions of manager and store fixed effects have about the same range in both companies, suggesting that changing a manager can have a productivity effect just as large as a hypothetical change in the store's otherwise persistent characteristics (assuming for the sake of argument that the observed relationships between store characteristics and productivity are causal).[16] In Company A, for example, replacing a very low-quality manager by a very high-quality one could increase the store's productivity by almost 45 log-points (over 50%). In Company B, the change would be considerably larger, though there is also more variation in the store

[15] We refer the reader to online Appendix C in Chandra et al. (2016) and the references therein for more details on the methodology employed.

[16] This observation might seem at odds with the conclusions from Table 1. However, note that the variances in that table are computed across all periods in the dataset, so the fixed effect of each store and manager is effectively weighted by the number of periods they are present in the data. The distributions in Figure 4 count each store/manager only once.

effects in this company as well.

Figure 4: Distribution of Manager and Store Quality (EB-Adjusted)



(a) Manager FE - Company A

(b) Store FE - Company A

(c) Manager FE - Company B

(d) Store FE - Company B

Figures A.2 and A.3 show that those findings are more broadly valid by plotting the EB-adjusted manager and store FE distributions in each of the four largest connected sets in our data. We note the caveat that the comparisons here are only valid inside connected sets, and it may be the case that there is more variation among stores and managers across connected sets than within them. In other words, managers' ties to store-level productivity might be even larger than the numbers we estimate here.

4 The Causes and Effects of Managerial Changes

We now explore what our data can say about the potential causes of a manager leaving a store, and what happens when a new manager is brought in. We estimate event study specifications assessing how store outcomes shift in the periods around a management change. We also ask whether stores that receive a new manager receive one who has a higher fixed effect than the manager who left, and similarly whether managers who switch stores end up moving to stores with higher fixed effects.

We estimate the following equation:

$$y_{st} = \alpha_s + \alpha_t + \mathbf{1}\{s \in \mathbb{T}\}\left[\sum_{k=-6}^{k=3} \beta_k \mathbf{1}\{K_{st} = k\} + \beta_{4+}\mathbf{1}\{K_{st} \geq 4\}\right] + \varepsilon_{st}, \tag{3}$$

where $y_{s,t}$ a store outcome; α_s and α_t are respectively store and time fixed effects; $\mathbf{1}\{s \in \mathbb{T}\}$ indicates whether store s belongs to the treated group $\mathbb{T}$ (i.e., changed its manager); and $K_{st} = t - E_s$ is the number of months before/after the time of event E_s (defined as the month when the outgoing manager leaves store s). Note that $\mathbf{1}\{s \in \mathbb{T}\}$ does not have a time subscript, indicating that stores that switch managers are in the treatment group even before the switch has happened. Stores in the control group are the ones that never change their manager. We define the treatment and control groups in this way for two important reasons. First, it avoids the identification issues discussed in Borusyak and Jaravel (2016). Second, a store manager change is often an expected event. It can therefore affect store outcomes before it actually happens, making simple before-and-after comparisons less informative.

We construct the control group for each treated store by finding the untreated stores that most resemble the treated store before treatment, according to the following metric: we estimate the conditional probability of treatment (propensity score) for each store s using observations dating to at least 6 months before the change in management takes place. Our empirical model is

$$\text{logit } \mathbb{P}(\text{mng change}_s) = \gamma_1 \log(revenue_s) + \gamma_2 \log(employment_s) + \gamma_3 \log(area_s) + X_s'\Gamma_0.$$

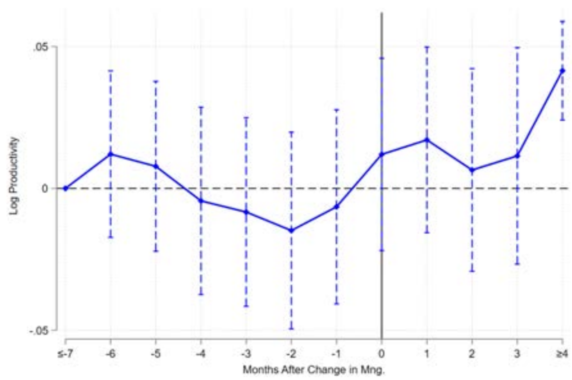
The dependent variable indicates if store s switched managers at any point in the data. The first set of regressors (sales revenue, FTE employment, and total floor area) control for a store's size and productivity. The remaining regressors in X_s control for the store's location (a state fixed effect) and entry and exit dates. All the right hand side variables are constructed by averaging the data for each store over time.

We match each treated store to its nearest untreated neighbors using the odds ratio as our distance metric.[17] We then construct the outcome of each control store $y_{s,t}(control)$ as the average of the outcomes of the 3 nearest neighbors of treated store s. Once this is done, we estimate (3) by noting that it maps into a potential outcomes framework

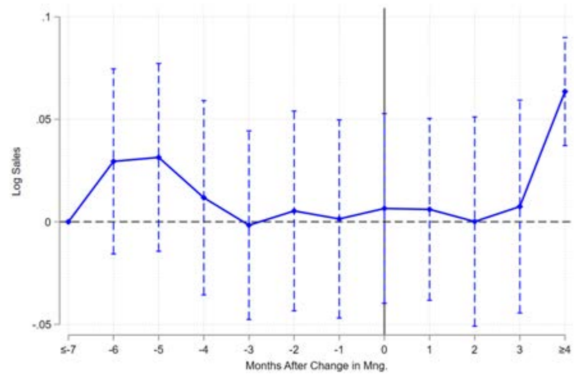
$$\Delta y_{st} = \sum_{k=-6}^{k=3} \beta_k \mathbf{1}\{K_{st} = k\} + \beta_{4+}\mathbf{1}\{K_{st} \geq 4\} + e_{st}, \tag{3'}$$

where $\Delta y_{st} = y_{s,t}(treated) - y_{s,t}(control)$. As discussed above, $y_{s,t}(treated)$ is the outcome of the treated store (observed) and $y_{s,t}(control)$ is the outcome of the constructed control for s.

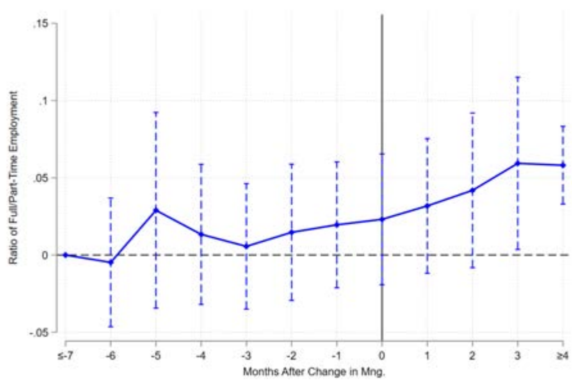
Figure 5: Event Studies: Changes in Outcomes Around Manager Change
Company A



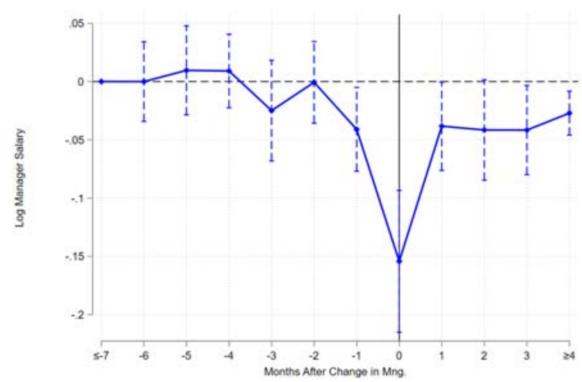
(a) Productivity



(b) Sales



(c) Full/Part-Time Emp. Ratio

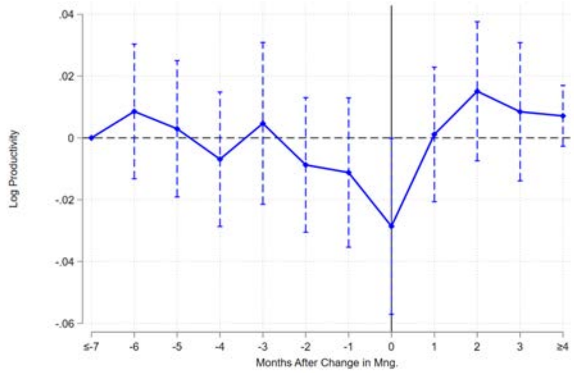


(d) Manager Salary

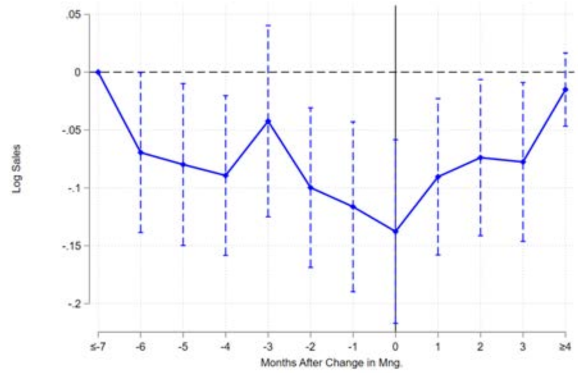
Note: Dots show the value of each coefficient β_k, while whiskers indicate the 95% confidence interval (computed using robust standard errors). The normalized period (≤ -7) indicates coefficients for periods that are precede the manager change by 7 or more months. Alternative specifications (e.g., using only 7-10 months prior to change) produce very similar results.

[17] Estimation is done through the Stata module psmatch2 (see Leuven and Sianesi, 2003).

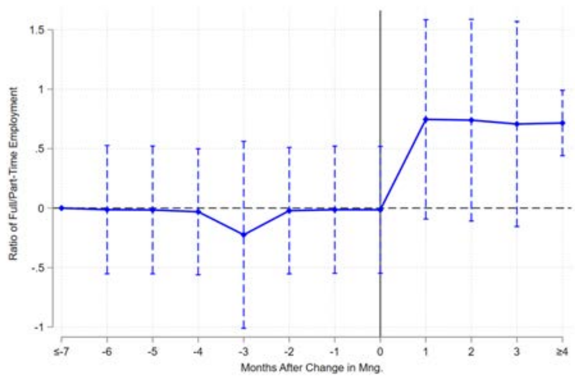
Figure 6: Event Studies: Changes in Outcomes Around Manager Change
Company B



(a) Productivity



(b) Sales



(c) Full/Part-Time Emp. Ratio

Note: Dots show the value of each coefficient β_k, while whiskers indicate the 95% confidence interval computed using robust standard errors. The normalized period (≤ -7) indicates coefficients for periods preceding the manager change by 7 or more months. Alternative specifications (e.g., using only 7-10 months prior to the manager change) produce similar results.

Figures 5 and 6 plot the estimated coefficients from model (3') for four outcome variables: productivity, sales, full-time to part-time employment ratio, and manager earnings (only available for Company A). Interestingly, the scenarios that precede a manager separation are very different across Companies A and B. This may highlight the importance of company culture or policy in determining manager mobility (and potentially impact).

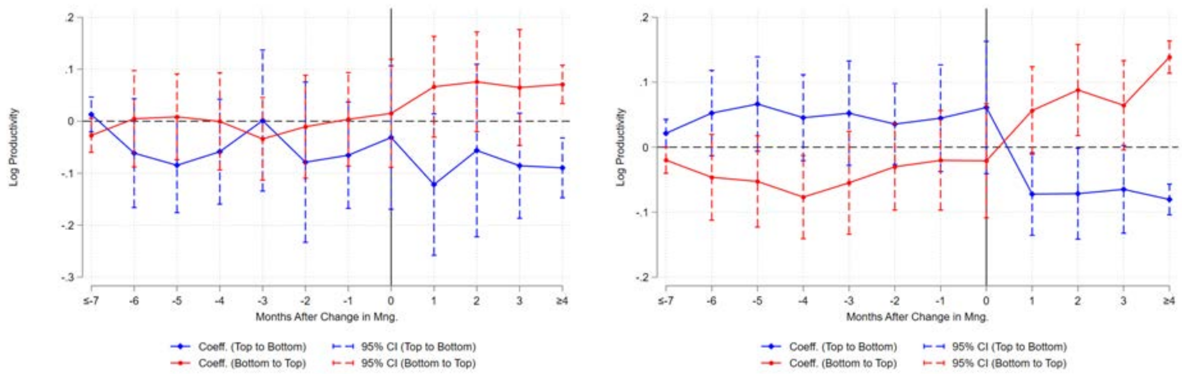
The event studies for Company A reveal no statistically significant change in productivity, sales, or employment composition around a managerial changeover. However, there is a large reduction in the managers' earnings (almost 30%) a couple of months before they leave their stores, and an immediate recovery once the new managers take over.

While manager salaries are not available for Company B, patterns of productivity, sales, and employment composition around a managerial change paint a completely different picture than seen in Company A. There is a steady decline in sales preceding a manager's departure, which causes a simultaneous decline in productivity. This trend reverts shortly after the arrival of the new manager, suggesting that original manager was removed after a few months of fading store performance. One pattern common to both companies is that the point estimates imply new managers accompany an increase in the ratio of full-time to part-time workers, though the standard errors are large enough that the change cannot be distinguished from zero at usual confidence levels. We refer to Appendix E for evidence on the absence of pre-trends during the months used for matching.

These patterns could be related to the reason for the manager's departure in the first place, say if bonuses drop due to poor performance. In Appendix E, we test whether the patterns are different for managers who leave the company altogether after moving from a store (excluding retirements) than for those who switch stores within the company. Managers remaining within the company may have different motivations for switching stores than managers who leave. We find that stayers have better sales and productivity performance than leavers, and the drop in their earnings right before switching stores is considerably attenuated (see figures E.3 and E.4). This is consistent with the hypothesis that some of the leavers may have been terminated or counselled out of the company. Interestingly, though, we find no difference in the distribution of the μ_m among the managers who leave compared to those who stay, suggesting that those ushered out of the company may have suffered performance declines that were out of their control.

We also study how changes from good to bad managers (and vice versa) affect stores in Figure E.5 and Appendix E. We group managers into top or bottom halves based on their position relative to the connected set median. We then repeat the event study analyses above for two groups of stores: those that start with a top manager and switch to a bottom manager, and those that start with a bottom manager and switch to a top manager. Our findings show an immediate impact of managers on store level productivity, with stores that switch from bottom to top managers experiencing a positive and permanent increase, and stores that switch from top to bottom managers experiencing the opposite. This is particularly evident in Company B (see Figure 7), but in both cases we see a productivity shift that is immediate and grows over time. In Company A, the earnings drop of top managers right before their move is also significantly less pronounced than the same among bottom managers (see Figures E.5 and E.6).[18]

Figure 7: Event Study: Switching Bottom and Top Managers



(a) Productivity (Company A) (b) Productivity (Company B)

Note: Dots show the value of each coefficient β_k, while whiskers indicate the 95% confidence interval (computed using robust standard errors). The red line represents coefficients for stores that switch from bottom (below median) to top (above median) managers; the blue line represents stores that switch from top to bottom managers. See Appendix E for details and the impacts on other select variables.

4.1 Are Stores Switching to Better Managers?

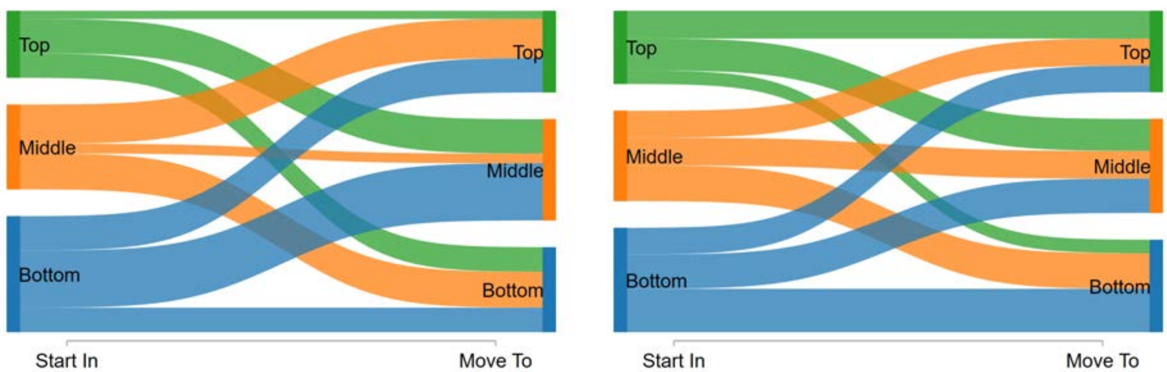
The event studies shed some light on what might lead to separations between managers and stores. We now investigate whether newly hired managers are, on average, different from those they replaced. We consider two manager attributes: their tenure in the company and their fixed effect. In both companies, the average incoming manager is less experienced than the outgoing manager. The average difference in the tenure of the outgoing and the incoming manager in Company A is 16.8 months, and significantly different from zero at the 1% level. In Company B the difference is also positive, but smaller: 6.5 months and not statistically significant at the usual levels. Compared to the range of the distribution of changes in tenure, those average differences are modest (see Figure A.6).

Figure A.7 shows the distribution of the differences in incoming and outgoing managers' EB-adjusted fixed effects. Panel (a) plots the results for Company A, and panel (b) does so for Company B. Because managers who have worked in the same store always share a connected set, their fixed effects can be directly compared. In both companies,

[18]All of these analyses and robustness checks bolster the case that many manager movements are not endogenously responding to expected performance changes, as sometimes has been argued in the CEO literature (Black, 2019).

the distribution is approximately symmetric around zero, with a mean difference not statistically different from zero at a 10% significance level. Thus, despite the circumstances that might lead a manager to leave a store, on average, the new manager is no different from the one they replaced.

Figure 8: Managerial Moves: Initial and Final Store Quality



(a) Company A (b) Company B

This figure shows the quality of the initial and final store where a manager worked, split into terciles based on the store's position in the productivity distribution. The size of each link in the chart is proportional to the share of flows it represents. Note that while stores are split into equally sized groups, the chart represents the number of *moves* from one group to another. Hence, if multiple managers move to/from the same store, that store will feature multiple times in the figure. As an example, many managers start in bottom stores, but not many managers move to bottom bottom stores, causing the difference in the proportion of bottom stores in each end of the chart.

The same exercise can be done from the viewpoints of managers: after moving to a different store, how does the new store quality compare to that of the one they left? Panels (c) and (d) of Figure A.7 show, for Companies A and B respectively, the distribution of changes in store fixed effects when managers move. On average, the store quality that a manager moves to increases in both companies, but is more pronounced in Company A. Not only is the average difference larger (approximately 0.05 for Company A versus 0.02 for Company B), the entire right side of Company A's distribution is bulkier. Figure 8, which plots the quality of the initial and final store where a manager works, demonstrates this result in Company A is mostly driven by managers who start in a store in the bottom of the productivity distribution and move to better stores. The equivalent flow is less pronounced in Company B.

That stores do not on average receive better managers, while managers move to better stores on average, might indicate interesting features of the manager-store allocation process for the companies in our data. One potential explanation is that manager quality is hard for the firms to measure and therefore to base allocative decisions on. Additionally, managers might have more incentive to learn store quality, as it affects their store's overall productivity and as such might affect bonuses or other types of compensation. We explore those possibilities in the next section. By simple comparison of means, we can say that managers in Company A tend to move to stores that (a) are larger, (b) pay higher wages, and (albeit with less statistical precision) (c) are located in more densely populated regions.

5 What Explains Manager and Store Productivity?

In this section, we explore whether observable characteristics and attributes in our data can explain variation in manager or store fixed effects. We do not view this exercise as pinning down causal relationships, but rather that the relationships among the variables might offer some understanding of the possible mechanisms at work in equilibrium store-manager allocations.

As a starting point, we assume that the quality of a manager can be described by

$$q_{m,t} = X_{s(m),t}\beta + q_{s(m),t} + \varepsilon_{m,t} \tag{4}$$

where $s(m), t$ indicates the store s where manager m worked in month t. Given our previous results, we also assume

$$q_{m,t} = \mu_m + u_{m,t} \quad \text{and} \quad q_{s,t} = \mu_s + u_{s,t}$$

where μ_m and μ_s are the manager and store fixed effects in equation (1). Plugging this expression into the equation above and taking the average over all periods (as the manager fixed effect is time-invariant), we find

$$\mu_m = \bar{X}_m\beta + \frac{1}{T}\sum_{t=1}^{T}\sum_{s\in S(m)} \mathbf{1}\{s = s(m), t\}\mu_s + \tilde{\varepsilon}_m,$$

where $S(m)$ is the set of stores where manager m was employed. Note that the right-hand side of this equation simply reflects the average store features ($\bar{X}_m$) and quality (μ_s) across each store where manager m was employed, weighted by the number of periods the manager spent at each store. In similar fashion, $\tilde{\varepsilon}_m$ is a combined residual that aggregates $\varepsilon_{m,t}$, $u_{m,t}$ and $u_{s(m),t}$.

Similarly, we have that the quality of stores can be expressed as

$$\mu_s = \bar{X}_s \beta + \frac{1}{T} \sum_{t=1}^{T} \sum_{m \in M(s)} \mathbf{1}\{m = m(s), t\} \mu_m + \tilde{\varepsilon}_s.$$

In our empirical analysis, we use the EB-adjusted manager and store fixed effects as the measure of μ_m and μ_s. To account for the fact that the fixed effects are normalized differently within each connected set, we include a full vector of connected set dummies in all regressions.

Manager/Store Quality and Consumer Satisfaction. For Company A, along with accounting data we also obtained the results from store-level consumer satisfaction surveys, starting in January 2019. These measure both overall satisfaction as well as impressions of more specific elements of the shopping experience.[19] As with productivity, we model the outcome of the consumer satisfaction survey as affected by both store and manager features. Hence, these outcomes are given by

$$CS_{s(m),t} = \delta_m + \delta_s + \delta_t + u_{s(m),t},$$

where $CS_{s(m),t}$ is the outcome of the survey for store s, where manager m was employed at time t. The δs represent fixed effects, and $u_{s,t}$ is a residual.

This specification presents two significant hurdles. First, it is not immediately clear which of the several possible outcomes/questions should be used to explain the variation in store/manager quality. Therefore we first calculate the first principal component of the scores for all outcomes, $CS^{(1PC)}_{s(m),t}$, and use that as a single-dimensional summary of consumer satisfaction.

Second, the estimation of δ_m and δ_s carry the same restrictions discussed in previous sections. In particular, the estimated δs are only comparable inside connected sets. However, because the consumer survey data is available only after January 2019, the network structure of this data (i.e., the connected sets) is different from the network structure of the accounting data, which starts in April 2018. This means that fixed effects estimated from the consumer survey data might not be comparable to the fixed effects estimated us-

[19] The full list includes: how likely the customer is to recommend the store, satisfaction with service received, satisfaction with the shopping experience, helpfulness, knowledgeable, being treated like a valued customer, availability, ease of checkout, appearance of the store, selection, value for price paid, and quality. Each of those is rated between 0 and 10, with 10 indicating the best outcome.

ing productivity. To avoid this issue, we instead regress $CS^{(1PC)}_{s(m),t}$ on a store (manager) and time fixed effect. The residual, $v_{m,t} = \delta_m + u_{s(m),t}$ (or $v_{s,t} = \delta_s + u_{s(m),t}$), is used as a measure of the manager's (store's) impact on consumer satisfaction. Once $v_{m,t}$ is computed, it can be inserted into (4), and we proceed the analysis in the same way as above.

The methods we use to deal with these issues may have important consequences for our results. We discuss alternatives and their outcomes in the robustness checks subsection below.

5.1 Results

Manager quality. Most observable features appear uncorrelated with managerial quality. Table 2 shows selected coefficients, with column (1) being the baseline regression, column (2) including the average EB-adjusted fixed effect across the stores where each manager has worked, and column (3) adding the consumer satisfaction survey outcomes (available only for company A).

The first thing to note from Table 2 is that there is no statistically significant relationship between a manager's quality and their salary. It is interesting to note that the coefficient on salary is negative in the baseline regression, but switches sign when the average store fixed effect is introduced. This could indicate that a large portion of a manager's compensation is due to the store where he/she works (note from Table 3 that better stores pay higher wages). However, the effect is too imprecise to draw strong conclusions.

Manager earnings are unfortunately not observed for Company B. However, we do observe stores' expenditures on energy inputs (electricity and gas). We construct a measure of energy efficiency as the ratio of sales revenue to energy costs and find that it is positively correlated with the manager fixed effect. To some extent, this result is expected. Energy is an input, and more productive managers would take actions to reduce costs per unit of output. In practice, however, this might not be as obvious, as energy efficiency tends not to be a focus of most retail managers. We develop this line of inquiry a bit more in Appendix F, where we present evidence that more productive managers reduce the consumption of energy itself, so that the perceived higher energy efficiency does not simply reflect lower energy prices. This is consistent with prior work by Bloom et al. (2010); Martin et al. (2012); Boyd and Curtis (2014), who find that well-managed firms have superior environmental performance. Our results suggest that managers themselves, independent of management policies and other company-level factors, influence energy performance.[20]

We also find that the ratio of full-time to part-time employees has an economically and statistically significant effect on the quality of managers. However, the sign of the coefficient is different across the two companies. This contrast result highlights a potential trade-off in hiring decisions: part-time workers offer more staffing flexibility, while full-time workers tend to be better trained and might reduce moral hazard issues. Company-wide features, as well as manager tastes, could influence which option is more appropriate. The number of hours that the store is open during weekdays is positively correlated with the manager's quality; this is expected as stores that are open for longer will mechanically have higher sales. However, the impact seems to be at least in part mediated through the match between manager and store store, as the coefficient on the log of hours open switches sign once we control for store quality. We should also note that, while manager plausibly have a say on the number of hours that a store is open, we do not see any changes in hours of operation in our sample.

Corroborating our earlier results, there is a negative relation between store and manager EB-adjusted fixed effects. Given its size and impact on other coefficients (e.g., comparing the last two columns for each company), the match between managers and stores appears to be one of the most important predictors of managerial quality (see also robustness checks below). Finally, we find a very small and statistically insignificant coefficient for the relationship between manager quality and consumer surveys. While puzzling, this is a robust result that we discuss more below.

It is also worth drawing attention to the R^2 of each regression. Very large portions of manager quality variation remain unexplained.[21] This suggests that potentially many other unobserved factors interact to determine managers' influence on productivity.

Store quality. We perform a similar exercise for stores. The results are in Table 3. Column (1) regresses the store's EB-adjusted fixed effect on attributes of the store's location,

[20] Managers contribute to their firm's energy productivity, so it is reasonable to think that managers can help explain the productivity dispersion within industries, as found in Lyubich, Shapiro and Walker (2018).

[21] Note that this is the *within-group* R^2; that is, it measures the explanatory power of the model *inside* each connected set. The overall R^2 in all regressions is above 0.99, but note that a lot of the variation in manager fixed effects across connected sets is in some sense artificial, as it is created by the normalization of store/manager effects within each connected set.

Table 2: Regression of EB-Adjusted Manager Fixed Effects on Observable Characteristics

	Company A			Company B	
	(1)	(2)	(3)	(1)	(2)
log manager salary	-0.045	0.011	0.042		
	(0.031)	(0.035)	(0.037)		
log energy efficiency				0.021	0.079***
				(0.029)	(0.020)
log FT/PT emp. ratio	0.058**	0.094***	0.098***	-0.026***	-0.044***
	(0.024)	(0.027)	(0.029)	(0.010)	(0.006)
log hours open: week	-0.079	0.537***	0.495**	0.039***	0.169***
	(0.239)	(0.200)	(0.208)	(0.093)	(0.059)
log hours open: weekend	-0.247	0.151	0.131	-0.002	0.007
	(0.327)	(0.237)	(0.233)	(0.051)	(0.022)
EB-adjusted store FE		-0.531***	-0.542***		-0.727***
		(0.080)	(0.084)		(0.060)
Consumer satisfaction (1PC)			-0.003		
			(0.005)		
N	1,222	738	624	2,120	1,541
Within R^2	0.025	0.334	0.363	0.012	0.587

Observations are weighted by the number of periods each manager is in the data. Standard errors are clustered at the connected set level and shown in parenthesis. ***, ** and * indicate that coefficients are significantly different from zero at the 1%, 5% and 10% levels, respectively.
Note: omitted controls include the number of hours open in the week and over the weekend, and the distance to each store's closest competitor.

namely population density and income per capita.[22] Column (2) adds the store's average manager compensation and full-time to part-time employment ratio. Column (3) adds the average EB-adjusted fixed effect of all managers who worked at the store, plus a location fixed effect to capture potential geographical features other than income and population density. Finally, column (4) includes the residualized first principal component of the consumer satisfaction survey, available only for company A. As before, all regressions also include a fixed effect for each connected set. For company B, we also include a dummy for the store's format (flagship store, small local, airport, etc).

More productive stores tend to be located in places with higher average income, but, taken at face value, our results suggest that there is no significant correlation with popula-

[22]For Company A, these values are for the store's county and are obtained through IPUMS-USA (Ruggles et al., 2021). For Company B, data are at the store's NUTS3-level region and are from the UK Office of National Statistics. Local income is measured by average wages in the American Community Survey (Company A), and by the regional gross disposable household income (Company B).

Table 3: Regression of EB-Adjusted Store Fixed Effects on Observable Characteristics

	Company A				Company B		
	(1)	(2)	(3)	(4)	(1)	(2)	(3)
log population density	0.025	0.005			-0.011	-0.010	
	(0.019)	(0.175)			(0.013)	(0.015)	
log income per capita	0.262**	-0.013			0.088	0.036	
	(0.102)	(0.094)			(0.118)	(0.119)	
log manager salary		0.446***	0.188	0.158			
		(0.120)	(0.148)	(0.151)			
log energy efficiency						0.149***	0.148***
						(0.031)	(0.018)
log FT/PT Emp. Ratio		0.141***	0.129**	0.141***		-0.060***	-0.059***
		(0.041)	(0.052)	(0.052)		(0.016)	(0.012)
log hours open: week		0.647**	1.076***	1.079***	0.238***	0.238***	
		(0.278)	(0.256)	(0.257)	(0.082)	(0.075)	
log hours open: weekend		0.895**	0.675*	0.741**	-0.101**	-0.057*	
		(0.438)	(0.346)	(0.337)	(0.041)	(0.031)	
EB-adjusted manager FE			-0.856***	-0.840***			-0.938***
			(0.194)	(0.198)			(0.027)
Consumer satisfaction (1PC)				-0.027			
				(0.020)			
Omitted controls	No	Yes	Yes	Yes	No	Yes	Yes
Location FE	No	No	Yes	Yes	No	No	Yes
N	278	276	229	229	919	826	880
Within R^2	0.043	0.374	0.588	0.595	0.023	0.113	0.576

Observations are weighted by the number of periods each store is in the data. Standard errors are clustered at the connected set level and shown in parenthesis. ***, ** and * indicate that coefficients are significantly different from zero at the 1%, 5% and 10% levels, respectively.
Note: omitted controls include: number of hours open in the week and over the weekend, distance to the closest competitor. Store size (log floor area) is also included as an omitted control for company A. All regressions for Company B include dummies for the store format (flagship, local store, etc.)

tion density. This is counter-intuitive given the well-studied prevalence of agglomeration economies and their impact on productivity (Duranton and Puga, 2004; Combes et al., 2012, among others). Indeed, this finding has one important caveat: because the regression includes connected set fixed effects, all identifying variation resides within those sets. If stores in the same connected set are also located in similar places, our coefficients understate the actual correlation between store quality and local features across the full sample. Appendix G explores the relationship between measures of agglomeration and store-level productivity in more detail, presenting novel evidence that agglomeration effects can be observed even *within firms*, not just across firms or cities.

We also find that high quality stores have managers who earn more, regardless of their quality. However, when the average quality of the managers in each store is included as an explanatory variable, the point estimate of the managerial salary coefficient drops by more than half and becomes statistically indistinguishable from zero. The coefficients on energy efficiency, full-time to part-time employment ratio, number of hours open, and average manager quality all follow the same pattern as in Table 2. Those results could just reflect similar underlying economic forces, or the fact that those attributes are influenced by both the manager and the store (e.g., through company policy or location).

5.2 Robustness Checks

We estimate several alternative models to asseses the robustness of our findings above. Our robustness checks can be roughly classified into three different categories, based on possible issues that arise in estimating the effect of manager or store characteristics in their quality.

Omitted variables. There are several potential features in a manager or store that might affect their impact on productivity. In our results above, we have focused on the ones that we consistently observe across both companies in our data. Among other observed characteristics that are either only available for a select number of managers or for only one company, we have manager tenure (both in the company and in the current store) and gender (only available for Company B). We do not find statistically significant relationships between manager tenure (or log-tenure, or tenure/log-tenure squared) or gender and the manager's FE.

Alternative measurements. Instead of including only the first principal component of the outcomes in consumer satisfaction surveys as a regressor in each regression, we can also include each individual outcome. Our results are not affected by this.

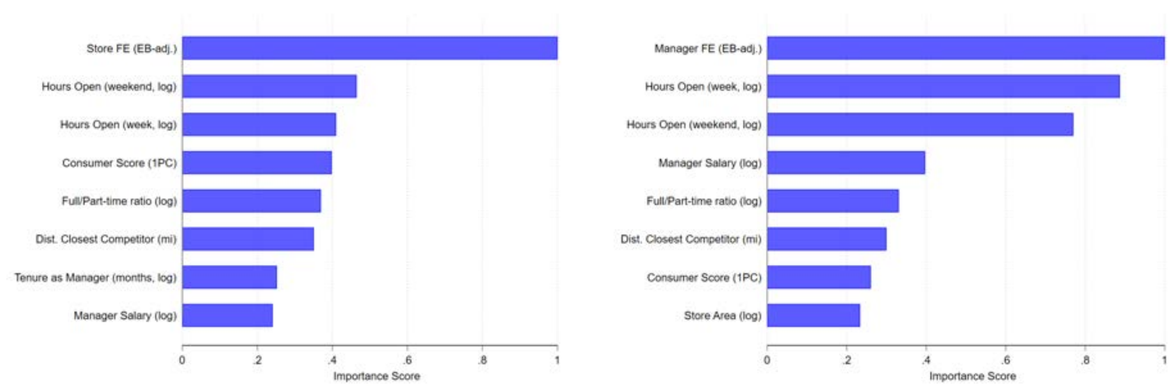
Alternatively, one could argue that our strategy of regressing $CS^{(1PC)}_{s(m),t}$ on time and store (or manager) fixed effects and using the residual $v_{m,t} = \delta_m + u_{s(m),t}$ as the contribution of the manager to the consumer satisfaction is prone to attenuation bias because this also includes the "pure" residual from the original regression. To address this concern, we can also directly estimate a manager fixed effect δ_m in

$$CS_{s(m),t} = X_{s,t}\beta + \delta_m + \delta_t + u_{s,t},$$

where $X_{s,t}$ are store-level characteristics meant to approximate the store's contribution to consumer satisfaction (including location, store size, number of employees, number of hours open, and distance to the closest competitor). The resulting relationship between δ_m and the manager's fixed effect is very similar to that found above.

Non-linear relationships. Finally, another possibility is that the linear model simply is not appropriate to explain the variation in the productivity of managers or stores. For that reason, we also estimate a random forest model that leverages the same variables shown in tables 2 and 3 to predict manager/store productivity. To account for the normalization issues across connected sets, we first de-mean all variables by their respective connected set (equivalent to including a connected set fixed effect in the linear model). Using the predicted values from the random forest model, we can calculate an R^2 statistic and compare its predictive power with the linear model.

Figure 9: Importance Matrix from Random Forest Model to Predict Manager and Store EB-Adjusted FE (Company A)



(a) Manager Productivity

(b) Store Productivity

Note: highest importance score normalized to 1 in each model.

As expected, the random forest model does a substantially better job at predicting manager and store productivity than the linear model does (in company A, R^2 increases to approximately 0.87 and 0.91 when explaining the variance in manager and store productivity, respectively).[23] However, impact of each explanatory follows the same pattern as in the linear model, which we can see by plotting the importance scores in Figure 9 (for brevity, we show only the results for company A).

[23]For this exercise, we run the random forest in the full dataset (there is no separate training data), so the exercise is comparable to the linear regression.

6 Previous Shocks and Adaptation to Covid-19

The Covid-19 pandemic severely disrupted life across the globe. Companies had to adapt quickly and substantially. In this section, we investigate whether good managers were able to better adapt to the Covid-19 pandemic restrictions, minimizing the potential negative effects on their store's performance. We begin by detailing some of the ways the pandemic affected managers and stores. Next, we investigate how the exposure to exogenous pre-Covid shocks (e.g., natural disasters, unsafe conditions, etc.) affected managers' ability to deal with the pandemic shock. Because of data availability, our results here apply only to Company A.

6.1 The Impact of Covid-19 Pandemic

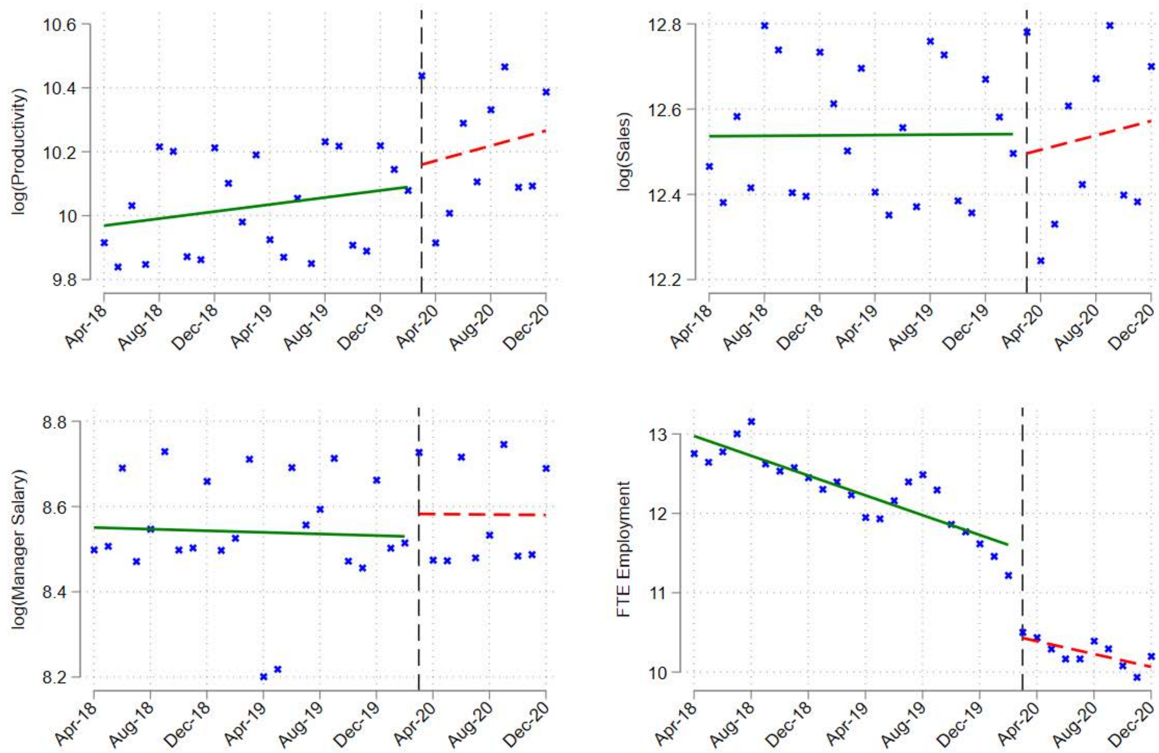
Data for Company A is available in a monthly basis between April 2018 and December 2020. We define the *Covid-19 period* as starting in March 2020, when the first non-pharmaceutical interventions (e.g., shelter-in-place and social distancing orders) were put in place in the US. This gives us 23 months of data before, and 10 months of data after the pandemic started.

Manager turnover: Compared to the pre-Covid period, manager turnover was slightly lower after the pandemic started. The probability that a store changed its manager in any given month was 3.4% before the pandemic and fell to 2.4% during the pandemic.

Store outcomes: Unlike what was experienced by many businesses, we do not observe a clear negative effect on stores' revenues or managers' earnings (see Figure 10). In contrast, there was a stark discontinuous decrease in full-time equivalent employment when the pandemic began. As a result, productivity at Company A's stores increased during Covid-19. This may reflect another instance of the "making do with less" effect (Lazear, Shaw and Stanton, 2016), sometimes observed during turbulent times or recessions.

Manager productivity: Managers who were more productive before the pandemic tend to perform better during the pandemic. We demonstrate this by testing the out-of-sample fit of equation (1). Specifically, we regress stores' log productivity levels on the EB-adjusted manager fixed effects, EB-adjusted store fixed effects, and a set of dummies for

Figure 10: Changes in Outcomes Since Covid-19



each month. Because the manager and store fixed effects were estimated in the pre-Covid period, the coefficient on the EB-adjusted fixed effects before Covid-19 is, by construction, one. When we fit the same model during the Covid-19 period, we find that the coefficient on the EB-adjusted manager fixed effect is a positive and significant (≈ 0.5), indicating that manager performance pre-pandemic is related to performance during the pandemic. Figure A.8 in the appendix plots the results.[24] The fact that the EB-adjusted manager FE coefficient is less than one in the latter period likely reflects a combination of mean reversion in manager performance (as suggested by Lazear, 2004), and the model's lower predictive power out of sample.

[24]We can also interpret this result as a out-of-sample test of our estimation of manager quality. To that end, Figure A.8 also performs this exercise with data from Company B, estimating the manager FEs in the first 30 months of data and measuring how those estimates explain variation in productivity in the latter 8 months in the data. In this case the coefficient is just over 0.85, indicating a a good out-of-sample fit. A plausible explanation for this coefficient being larger than Company A's is that Covid-19 induced substantial changes in retail operations.

6.2 Exposure to Shocks and Productivity During the Pandemic

The Covid-19 pandemic and the resulting containment measures and changes in consumer behavior can be interpreted as a series of exogenous shocks causing the temporary closure of stores, reductions in in-store customers and staff, employee absenteeism, new cleaning protocols, and so on. Such disruptions in daily operations had been experienced by many managers before Covid-19, although to a smaller degree. Examples include weather events like winter storms and hurricanes, power outages, police activity, and problems with stores' facilities. We have data for Company A on a large number and variety of such events both before and after the pandemic started (a full list of the shocks and their effects on store operations is shown in Table B.4). We use it here to see if managers' experiences with such unexpected challenges prior to Covid-19 influenced their performance during the pandemic.

We start by regressing log productivity during Covid-19 on the previously estimated manager and store EB-adjusted fixed effects, a set of variables recording the number of shocks a store experiences in a given month (one variable for each type of shock), calendar month dummies, and an indicator for each connected set. The coefficients are shown in columns (1) and (2) of Table 4. Most recorded shocks do not have a statistically significant effect on store productivity; for the sake of space, we only show the coefficient for "unsafe conditions due to Covid-19," which has by far the largest productivity impact.

A store deemed unsafe experiences a severe drop in monthly productivity. The results in column (1) imply an average drop of 43 log points, approximately 35%. In column (2) we allow the effect of unsafe conditions to vary with manager quality. We find a large base effect (for a manager at the mean of the FE distribution) of 57 log points, corresponding to a 43% drop in store productivity. However, better managers mitigate part of the shock's effect, at a rate of about 6 percentage points per "unit" of manager quality. To put this number into perspective, recall that the range of manager fixed effects in the largest connected set in Company A is about 0.5 (see Figure 4). This means that a top manager required to close their store due to unsafe conditions will experience a productivity decline 3 percentage points (0.5×6) smaller than a store whose manager is in the left tail of the productivity distribution.

Columns (3) and (4) of Table 4 show the results of a similar regression, but instead analyze the relationship between store productivity during Covid-19 and the total number of pre-pandemic shocks experienced by managers. Here, we include manager tenure as a

Table 4: Impact of Manager Experience with Previous Shocks on Store Productivity During Covid-19 (Company A)

	Effect of Shocks		Effect of Experience	
	(1)	(2)	(3)	(4)
Manager FE (pre-Covid; EB-Adj.)	0.503***	0.503***	0.498***	0.479***
	(0.043)	(0.043)	(0.045)	(0.046)
Store FE (pre-Covid; EB-Adj.)	0.649***	0.650***	0.649***	0.653***
	(0.028)	(0.028)	(0.029)	(0.029)
Unsafe conditions (due to Covid-19)	-0.434***	-0.567***		
	(0.129)	(0.144)		
Manager FE × Unsafe conditions		0.059***		
		(0.015)		
# Previous shocks			0.010*	0.019**
			(0.006)	(0.007)
Manager FE × # Previous shocks				-0.004***
				(0.001)
Experienced shock (current month)				-0.034***
				(0.011)
Experienced shock × # Previous shocks				-0.015
				(0.013)
N	3,108	3,108	3,108	3,108
Within R^2	0.2212	0.2370	0.1788	0.1878

Robust standard errors in parenthesis and ***, ** and * indicate coefficients are significantly different from zero at the 1%, 5% and 10% levels, respectively. All specifications include a fixed effect for each time period and connected set. Columns (1) and (2) also have dummies for all other observed shocks, with their interactions with manager fixed effects also included in column (2). Columns (3) and (4) also control for manager tenure.
Note: unlike in previous regressions, standard errors are not clustered by connected sets, as the connected sets before and after the pandemic are different.

control variable, as managers with longer tenure would mechanically have had more exposure to shocks. The results in column (3) indicate that managers' previous experiences with shocks are related to better store productivity during the pandemic, though this is imprecise. When we interact the number of previous shocks with managerial quality in column (4), we find a positive and significant main effect and a negative interaction. One possible interpretation is that worse managers gain more from previous experiences with shocks, potentially because they had more room to grow in the first place. The regression in column (4) also includes a dummy that indicates whether the store experienced any type of shock in the current month, as well as its interaction with the number of past shocks experienced by the manager. Unsurprisingly, shocks negatively affect store productivity during the pandemic period. We do not find evidence that experience with past shocks alone enables managers to deal with current shocks better than their peers.

6.2.1 Robustness Checks

Severity of shocks. Better managers appear to attenuate the productivity effect of a Covid-19 shock once it happens. We further explore the connections between manager quality and the severity of shocks by determining whether specific types of shocks have differential effects on stores operating practices (upstream of their observed productivity). We first rank the severity of shocks based on their influence on store operations. From least to most severe, they are:

1: Open late, mid-day closure, close early

2: Full day closure (one day only)

3: Temporarily reducing operating hours, switching to curbside

4: Long-term closure (3 or more days)

We regress this severity index on the manager and store fixed effects for those that experienced each shock, a set of indicators for each type of shock (second column of Table B.4), and indicators for each month and connected set. In an alternative specification, we also interact manager quality and the type of shock. Because of the ordinal nature of the index, we estimate these specifications as an ordered logit. We do not find a statistically significant relationship between manager (or store) FEs and the severity of shocks.

Exposure to shocks and the likelihood of leaving. Finally, we check if managers who are exposed to more shocks before the pandemic are also more likely to leave the company. If so, the results above linking managers' exposure to shocks and post-Covid performance might have a selection issue.

We estimate a linear probability model of whether a manager left Company A before February of 2020 on the number of shocks the manager was exposed to during the pre-pandemic period plus other store-level observables (table 5). Once again, because the variable of interest is fixed over time, all of the explanatory variables are first averaged/summed to facilitate estimation of a cross-sectional model.

Table 5: Probability of Leaving Company A in the Pre-pandemic Period

	(1)	(2)	(3)
Total # previous shocks	-0.002	-0.002	-0.005***
	(0.002)	(0.002)	(0.001)
Avg. log manager salary	-0.435***	-0.384**	-0.384**
	(0.164)	(0.180)	(0.185)
log manager tenure (last period)	0.110***	0.058**	0.062**
	(0.023)	(0.025)	(0.024)
Last store FE (EB-Adj.)		-0.073	-0.080
		(0.138)	(0.140)
Manager FE (EB-Adj.)		0.089	0.080
		(0.028)	(0.229)
Manager FE × # previous shocks			0.001***
			(0.0004)
N	545	443	443
Within R^2	0.0773	0.0326	0.0696

Robust standard errors in parenthesis and ***, ** and * indicate coefficients are significantly different from zero at the 1%, 5% and 10% levels, respectively. All specifications include a fixed effect for each connected set.

Unsurprisingly, managers with high salaries are less likely to leave the company, even after controlling for manager and store quality. In contrast, managers with long tenures are more likely to leave (departures include retirements). That said, we do not find any sign that managers with high exposure to shocks are more likely on average to leave Company A. There is a small directional interaction: the likelihood of a manager leaving as they experience more shocks increases slightly faster for high-FE than low-FE managers. However, this difference is incredibly small. Compared to a manager at the bottom of the productivity distribution that experienced no shocks, a top manager who experienced 10 shocks (the 90th percentile of the distribution) is only 0.5 percentage points more likely to leave.

7 Conclusion

A growing literature has established that management influences business performance and productivity. This work has spanned settings that vary both by industry/sector and the development level of the broader economy the business operates within.

In this paper we show that managers have an impact on productivity that is distinct, relatively quick, and separate from company-level management practices. Quantitatively, we find that establishment-level characteristics can explain a larger share of the variance in productivity than managers can, but that the ranges of the distributions of the two effects are comparable. In that sense, we estimate that replacing manager at the bottom of the distribution by one at the top could increase a store's productivity by at least 50%, and perhaps as much as doubling it, depending on the company and the relevant connected set.

There are some obvious caveats to our study. First, we only have data for two retail firms. While it is not particularly difficult to imagine at least some of our findings applying more generally, we cannot test this supposition. Second, while our ability to hold firm-level management practices fixed is an advantage that allows us to zoom in on the roles of managers as individuals, we cannot impose common practices within those firm-level bounds. Relatedly, we do not have access to granular data of what managers do on a moment-to-moment basis or a measure of their social and strategy skills. This limits our ability to understand the particular mechanisms that underlie our findings, given that such variables have been found to be correlated with productivity and satisfcation in the workplace (Antonakis et al., 2014; Bandiera et al., 2020; Hansen et al., 2021; Impink, Prat and Sadun, 2021; Dube, Naidu and Reich, 2022; Alan, Corekcioglu and Sutter, 2023). In this way, managers and practices are to some extent still inherently linked in our setting. Third, our sample period is relatively short, so we cannot fully characterize any dynamic effects, or estimate whether manager fixed effects experience low-frequency drifts due to learning or training.

These caveats aside, we believe our results point to there being much more fruitful work to be done in understanding how managers influence the productivity of the establishments they manage.

References

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica*, 67(2): 251–333.

Acemoglu, Daron, Alex He, and Daniel le Maire. 2022. "Eclipse of Rent-Sharing: The Effects of Managers' Business Education on Wages and the Labor Share in the US and Denmark." National Bureau of Economic Research.

Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham. 2019. "Management and shocks to worker productivity." National Bureau of Economic Research.

Adhvaryu, Achyuta, Vittorio Bassi, Anant Nyshadham, and Jorge A Tamayo. 2020. "No line left behind: Assortative matching inside the firm." National Bureau of Economic Research.

Aghion, Philippe, Nicholas Bloom, Brian Lucking, Raffaella Sadun, and John Van Reenen. 2021. "Turbulence, firm decentralization, and growth in bad times." *American Economic Journal: Applied Economics*, 13(1): 133–69.

Alan, Sule, Gozde Corekcioglu, and Matthias Sutter. 2023. "Improving workplace climate in large corporations: A clustered randomized intervention." *Quarterly Journal of Economics*, 138(1): 151–203.

Aldy, Joseph E, and William A Pizer. 2015. "The competitiveness impacts of climate change mitigation policies." *Journal of the Association of Environmental and Resource Economists*, 2(4): 565–595.

Andrews, Dan, Andrew Charlton, and Angus Moore. 2021. "COVID-19, Productivity and Reallocation: Timely evidence from three OECD countries."

Andrews, M. J., L. Gill, T. Schank, and R. Upward. 2008. "High wage workers and low wage firms: negative assortative matching or limited mobility bias?" *Journal of the Royal Statistical Society Series A*, 171(3): 673–697.

Antonakis, John, Giovanna d'Adda, Roberto Weber, and Christian Zehnder. 2014. "Just words? Just speeches? On the economic value of charismatic leadership." *NBER Rep*, 4.

Atkeson, Andrew, and Patrick J Kehoe. 1999. "Models of energy use: Putty-putty versus putty-clay." *American Economic Review*, 89(4): 1028–1043.

Auffhammer, Maximilian, and Erin T Mansur. 2014. "Measuring climatic impacts on energy consumption: A review of the empirical literature." *Energy Economics*, 46: 522–530.

Baltrunaite, Audinga, Giulia Bovini, and Sauro Mocetti. 2023. "Managerial talent and managerial practices: are they complements?" *Journal of Corporate Finance*, 102348.

Bandiera, Oriana, Andrea Prat, Stephen Hansen, and Raffaella Sadun. 2020. "CEO behavior and firm performance." *Journal of Political Economy*, 128(4): 1325–1369.

Bandiera, Oriana, Luigi Guiso, Andrea Prat, and Raffaella Sadun. 2015. "Matching firms, managers, and incentives." *Journal of Labor Economics*, 33(3): 623–681.

Bao, Renjie, Jan De Loecker, and Jan Eeckhout. 2022. "Are Managers Paid for Market Power?" National Bureau of Economic Research Working Paper 29918.

Barberis, Nicholas, Maxim Boycko, Andrei Shleifer, and Natalia Tsukanova. 1996. "How does privatization work? Evidence from the Russian shops." *Journal of Political Economy*, 104(4): 764–790.

Barry, John W, Murillo Campello, John Graham, and Yueran Ma. 2022. "Corporate flexibility in a time of crisis." National Bureau of Economic Research.

Becker, Gary S. 1973. "A theory of marriage: Part I." *Journal of Political Economy*, 81(4): 813–846.

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud. 2014. "Productive cities: Sorting, selection, and agglomeration." *Journal of Political Economy*, 122(3): 507–553.

Bell, Brian D, Nicholas Bloom, and Jack Blundell. 2021. "This Time is Not so Different: Income Dynamics During the COVID-19 Recession." National Bureau of Economic Research.

Bender, Stefan, Nicholas Bloom, David Card, John Van Reenen, and Stefanie Wolter. 2018. "Management practices, workforce selection, and productivity." *Journal of Labor Economics*, 36(S1): S371–S409.

Benson, Alan, Danielle Li, and Kelly Shue. 2019. "Promotions and the Peter principle." *Quarterly Journal of Economics*, 134(4): 2085–2134.

Berndt, Ernst R, and David O Wood. 1979. "Engineering and econometric interpretations of energy-capital complementarity." *American Economic Review*, 69(3): 342–354.

Bertrand, Marianne, and Antoinette Schoar. 2003. "Managing with Style: The Effect of Managers on Firm Policies." *Quarterly Journal of Economics*, 118(4): 1169–1208.

Best, Michael Carlos, Jonas Hjort, and David Szakonyi. 2017. "Individuals and organizations as sources of state effectiveness." National Bureau of Economic Research.

Bianchi, Nicola, and Michela Giorcelli. 2022. "The dynamics and spillovers of management interventions: Evidence from the training within industry program." *Journal of Political Economy*, 130(6): 1630–1675.

Black, Ines. 2019. "Better together? CEO identity and firm productivity." *CEO Identity and Firm Productivity (March 20, 2019)*.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices Across Firms and Countries." *Quarterly Journal of Economics*, 122(4): 1351–1408.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does Management Matter? Evidence from India." *Quarterly Journal of Economics*, 128(1): 1–51.

Bloom, Nicholas, Christos Genakos, Ralf Martin, and Raffaella Sadun. 2010. "Modern management: good for the environment or just hot air?" *Economic Journal*, 120(544): 551–572.

Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen. 2019. "What drives differences in management practices?" *American Economic Review*, 109(5): 1648–83.

Bloom, Nicholas, Philip Bunn, Paul Mizen, Pawel Smietanka, and Gregory Thwaites. 2020. "The impact of Covid-19 on productivity." National Bureau of Economic Research.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. 2012. "The organization of firms across countries." *Quarterly Journal of Economics*, 127(4): 1663–1705.

Bloom, Nicholas, Renata Lemos, Raffaella Sadun, Daniela Scur, and John Van Reenen. 2014. "The New Empirical Economics of Management." *Journal of the European Economic Association*, 12(4): 835–876.

Bloom, Nicholas, Robert S Fletcher, and Ethan Yeh. 2021. "The impact of COVID-19 on US firms." National Bureau of Economic Research.

Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler. 2020. "How Much Should we Trust Estimates of Firm Effects and Worker Sorting?" NBER Working Paper 27368.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa. 2019. "Discretizing Unobserved Heterogeneity." University of Chicago, Becker Friedman Institute for Economics.

Borusyak, Kirill, and Xavier Jaravel. 2016. "Revisiting Event Study Designs." Working Paper.

Boudreau, Laura, Rocco Macchiavello, Virginia Minni, and Mari Tanaka. 2021. "Union Leaders: Experimental Evidence From Myanmar."

Boyd, Gale A, and E Mark Curtis. 2014. "Evidence of an Energy-Management Gap in US manufacturing: Spillovers from firm management practices to energy efficiency." *Journal of Environmental Economics and Management*, 68(3): 463–479.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar. 2018. "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico." *Journal of Political Economy*, 126.

Cai, Jing, and Shing-Yi Wang. 2022. "Improving management through worker evaluations: Evidence from auto manufacturing." *Quarterly Journal of Economics*, 137(4): 2459–2497.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and labor market inequality: Evidence and some theory." *Journal of Labor Economics*, 36(S1): S13–S70.

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace heterogeneity and the rise of West German wage inequality." *Quarterly journal of economics*, 128(3): 967–1015.

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson. 2016. "Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector." *American Economic Review*, 106(8): 2110–44.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux. 2012. "The productivity advantages of large cities: Distinguishing agglomeration from firm selection." *Econometrica*, 80(6): 2543–2594.

Correia, Sergio. 2017. "Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator."

Cowgill, Bo, Jonathan Davis, B Pablo Montagnes, and Patryk Perkowski. 2021. "Matchmaking Principals: Theory and Evidence from Internal Labor Markets."

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum. 2022. "Matching in cities." *Journal of the European Economic Association*, 20(4): 1478–1521.

Davis, Steven J, Cheryl Grim, John Haltiwanger, and Mary Streitwieser. 2013. "Electricity unit value prices and purchase quantities: US manufacturing plants, 1963–2000." *Review of Economics and Statistics*, 95(4): 1150–1165.

Decker, Ryan A, John C Haltiwanger, Ron S Jarmin, and Javier Miranda. 2018. "Changing business dynamism and productivity: Shocks vs Responsiveness." responsiveness. Technical report, National Bureau of Economic Research.

DellaVigna, Stefano, and Matthew Gentzkow. 2019. "Uniform pricing in us retail chains." *Quarterly Journal of Economics*, 134(4): 2011–2084.

Deschênes, Olivier, and Michael Greenstone. 2011. "Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US." *American Economic Journal: Applied Economics*, 3(4): 152–85.

Di Addario, Sabrina, Patrick Kline, Raffaele Saggio, and Mikkel Sølvsten. 2022. "It ain't where you're from, it's where you're at: hiring origins, firm heterogeneity, and wages." *Journal of Econometrics*.

Dube, Arindrajit, Suresh Naidu, and Adam D Reich. 2022. "Power and Dignity in the Low-Wage Labor Market: Theory and Evidence from Wal-Mart Workers." National Bureau of Economic Research.

Duranton, Gilles, and Diego Puga. 2004. "Micro-foundations of urban agglomeration economies." In *Handbook of Regional and Urban Economics*. Vol. 4, 2063–2117. Elsevier.

Eeckhout, Jan, and Philipp Kircher. 2011. "Identifying sorting – in theory." *Review of Economic Studies*, 78(3): 872–906.

Eeckhout, Jan, and Philipp Kircher. 2018. "Assortative matching with large firms." *Econometrica*, 86(1): 85–132.

Fee, C Edward, Charles J Hadlock, and Joshua R Pierce. 2013. "Managers with and without style: Evidence using exogenous variation." *Review of Financial Studies*, 26(3): 567–601.

Fenizia, Alessandra. 2022. "Managers and productivity in the public sector." *Econometrica*, 90(3): 1063–1084.

Frederiksen, Anders, Lisa B Kahn, and Fabian Lange. 2020. "Supervisors and performance management systems." *Journal of Political Economy*, 128(6): 2123–2187.

Friebel, Guido, Matthias Heinz, and Nikolay Zubanov. 2022. "Middle Managers, Personnel Turnover, and Performance: A Long-Term Field Experiment in a Retail Chain." *Management Science*, 68(1): 211–229.

Gaure, Simen. 2013. "lfe: Linear group fixed effects." *The R Journal*, 5(2): 104–117. User documentation of the 'lfe' package.

Gaure, Simen. 2014. "Correlation bias correction in two-way fixed effects linear regression." *Stat*, 3(1): 379–390. Description of the limited mobility bias correction method used in 'lfe'.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer. 2013. "Human capital and regional development." *Quarterly Journal of Economics*, 128(1): 105–164.

Giardili, Soledad, Kamalini Ramdas, and Jonathan W Williams. 2022. "Leadership and productivity: a study of US automobile assembly plants." *Management Science*.

Giorcelli, Michela. 2019. "The Long-Term Effects of Management and Technology Transfers." *American Economic Review*, 109(1): 121–52.

Gosnell, Greer K., John List, and Robert Metcalfe. 2020. "The Impact of Management Practices on Employee Productivity: A Field Experiment with Airline Captains." *Journal of Political Economy*, 128(4): 1195 – 1233.

Hagedorn, Marcus, Tzuo Hann Law, and Iourii Manovskii. 2017. "Identifying equilibrium models of labor market sorting." *Econometrica*, 85(1): 29–65.

Hansen, Stephen, Tejas Ramdas, Raffaella Sadun, and Joe Fuller. 2021. "The Demand for Executive Skills." National Bureau of Economic Research.

Hoffman, Mitchell, and Steven Tadelis. 2021. "People management skills, employee attrition, and manager rewards: An empirical analysis." *Journal of Political Economy*, 129(1): 243–285.

Hoffman, Mitchell, Lisa B Kahn, and Danielle Li. 2018. "Discretion in hiring." *Quarterly Journal of Economics*, 133(2): 765–800.

Huber, Kilian, Volker Lindenthal, and Fabian Waldinger. 2021. "Discrimination, managers, and firm performance: Evidence from aryanizations in nazi Germany." *Journal of Political Economy*, 129(9): 2455–2503.

Ichniowski, Casey, and Kathryn Shaw. 1999. "The effects of human resource management systems on economic performance: An international comparison of US and Japanese plants." *Management science*, 45(5): 704–721.

Ichniowski, Casey, and Kathryn Shaw. 2003. "Beyond incentive pay: Insiders' estimates of the value of complementary human resource management practices." *Journal of Economic Perspectives*, 17(1): 155–180.

Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi. 1997. "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines." *American Economic Review*, 87(2): 768–809.

Impink, Stephen Michael, Andrea Prat, and Raffaella Sadun. 2021. "Communication within Firms: Evidence from CEO Turnovers." National Bureau of Economic Research.

Janke, Katharina, Carol Propper, and Raffaella Sadun. 2019. "The impact of CEOs in the public sector: Evidence from the English NHS." National Bureau of Economic Research.

Kahn, Matthew E, and Erin T Mansur. 2013. "Do local energy prices and regulation affect the geographic concentration of employment?" *Journal of Public Economics*, 101: 105–114.

Kahn, Matthew E., and Nils Kok. 2014. "Big-Box Retailers and Urban Carbon Emissions: The Case of Wal-Mart." National Bureau of Economic Research, Inc NBER Working Papers 19912.

Kaplan, Steven N, and Morten Sorensen. 2017. "Are CEOs different? Characteristics of top managers." National Bureau of Economic Research.

Lachowska, Marta, Alexandre Mas, Raffaele Saggio, and Stephen A Woodbury. 2022. "Do firm effects drift? Evidence from Washington administrative data." *Journal of Econometrics*.

Lazear, Edward P. 2004. "The Peter Principle: A theory of decline." *Journal of political economy*, 112(S1): S141–S163.

Lazear, Edward P, Kathryn L Shaw, and Christopher Stanton. 2016. "Making do with less: working harder during recessions." *Journal of Labor Economics*, 34(S1): S333–S360.

Lazear, Edward P, Kathryn L Shaw, and Christopher T Stanton. 2015. "The value of bosses." *Journal of Labor Economics*, 33(4): 823–861.

Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet. 2021. "Gender differences in job search: Trading off commute against wage." *Quarterly Journal of Economics*, 136(1): 381–426.

Leuven, Edwin, and Barbara Sianesi. 2003. "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing." *Statistical Software Components, Boston College Department of Economics*.

Limodio, Nicola. 2021. "Bureaucrat allocation in the public sector: evidence from the World Bank."

Lopes de Melo, Rafael. 2018. "Firm wage differentials and labor market sorting: Reconciling theory and evidence." *Journal of Political Economy*, 126(1): 313–346.

Lyubich, Eva, Joseph S Shapiro, and Reed Walker. 2018. "Regulating mismeasured pollution: Implications of firm heterogeneity for environmental policy." Vol. 108, 136–142, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Marin, Giovanni, and Francesco Vona. 2021. "The impact of energy prices on socioeconomic and environmental performance: Evidence from French manufacturing establishments, 1997–2015." *European Economic Review*, 135: 103739.

Martin, Ralf, Mirabelle Muûls, Laure B De Preux, and Ulrich J Wagner. 2012. "Anatomy of a paradox: Management practices, organizational structure and energy efficiency." *Journal of Environmental Economics and Management*, 63(2): 208–223.

Minni, V. 2022. "Making the Invisible Hand Visible: Managers and the Allocation of Workers to Jobs." mimeo.

Mollick, Ethan. 2012. "People and process, suits and innovators: The role of individuals in firm performance." *Strategic Management Journal*, 33(9): 1001–1015.

Patault, Berengere, and C Lenoir. 2020. "How valuable are business networks? evidence from sales managers in international markets." Working paper.

Pindyck, Robert S, and Julio J Rotemberg. 1983. "Dynamic factor demands and the effects of energy price shocks." *American Economic Review*, 73(5): 1066–1079.

Popp, David. 2002. "Induced innovation and energy prices." *American Economic Review*, 92(1): 160–180.

Rubens, Michael. 2022. "Management, productivity, and technology choices: Evidence from US mining schools." *RAND Journal of Economics*.

Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek. 2021. "IPUMS USA: Version 11.0 [dataset]." Minneapolis, MN: IPUMS.

Sandvik, Jason J, Richard E Saouma, Nathan T Seegert, and Christopher T Stanton. 2020. "Workplace knowledge flows." *Quarterly Journal of Economics*, 135(3): 1635–1680.

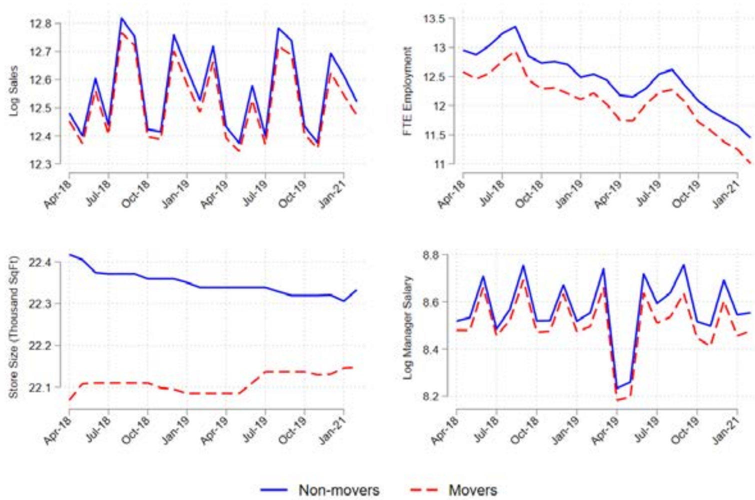
Singer, Gregor. 2022. "Complementary Inputs and Industrial Development: Can Lower Electricity Prices Improve Energy Efficiency?"

Sorkin, Isaac. 2018. "Ranking firms using revealed preference." *Quarterly Journal of Economics*, 133(3): 1331–1393.

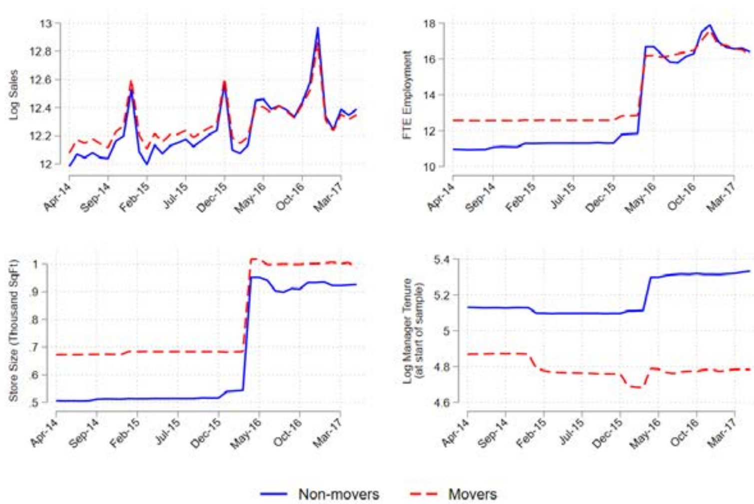
Syverson, Chad. 2011. "What Determines Productivity?" *Journal of Economic Literature*, 49(2): 326–65.

A Figures

Figure A.1: Pre-move Trends: Difference Between Movers and Non-movers



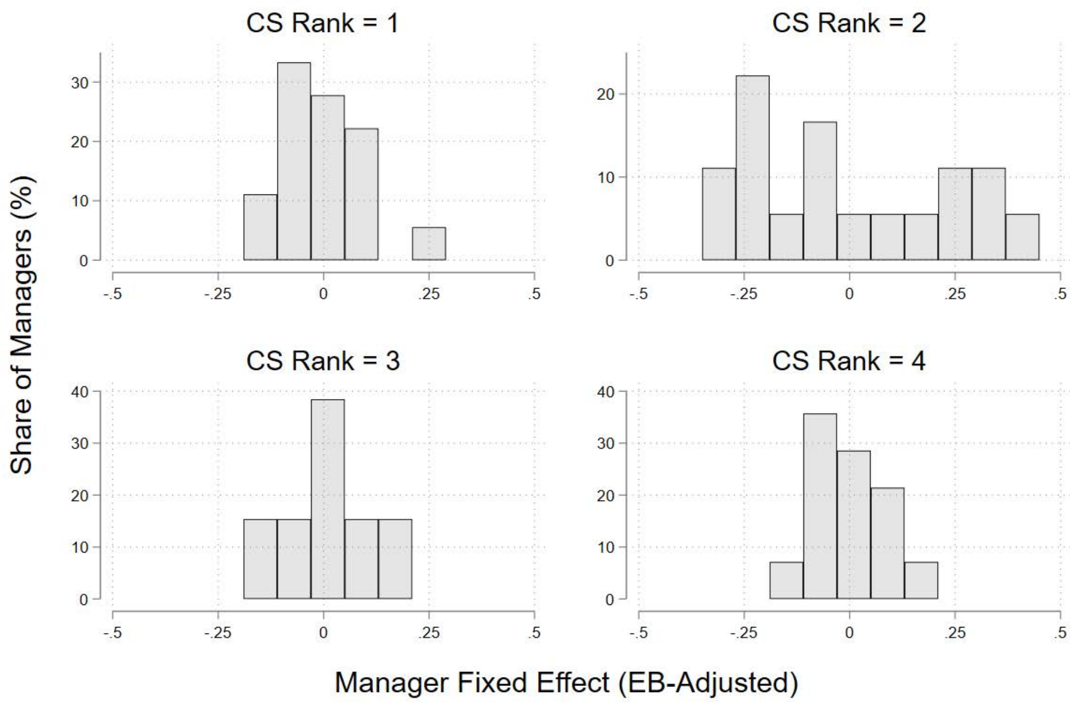
(a) Company A



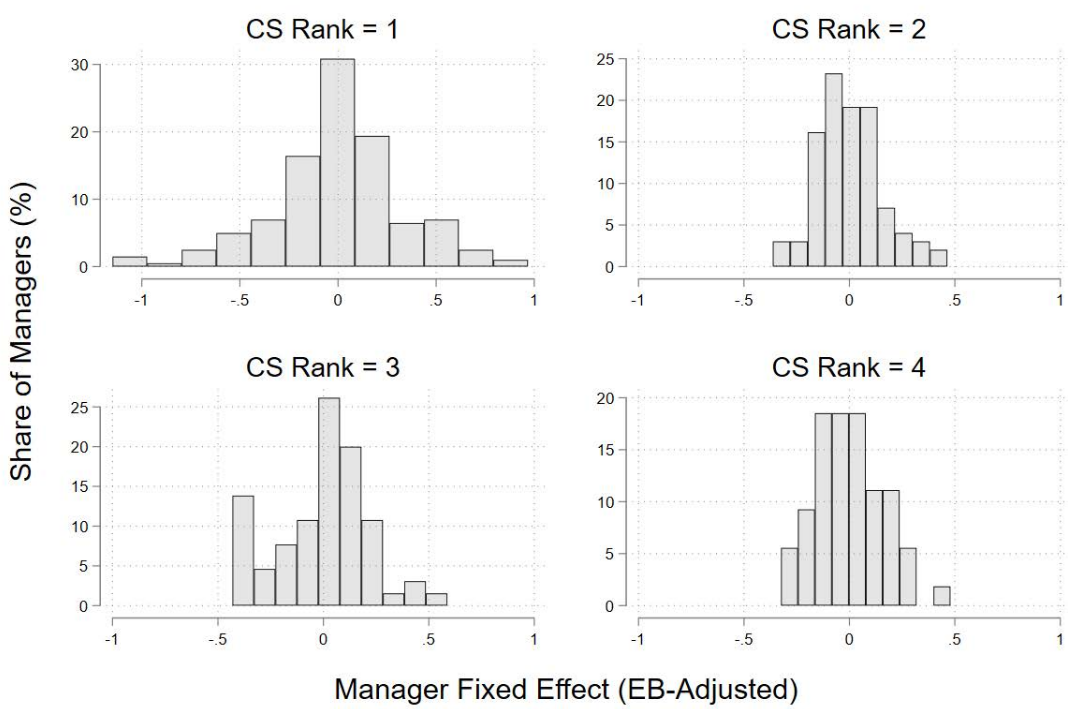
(b) Company B

Note: starting in April of 2016, data for many of the smaller establishments of Company B is no longer available, which accounts for the shifts in average outcomes (for both movers and non-movers) at that date.

Figure A.2: Distribution of EB-Adjusted Manager Fixed Effects



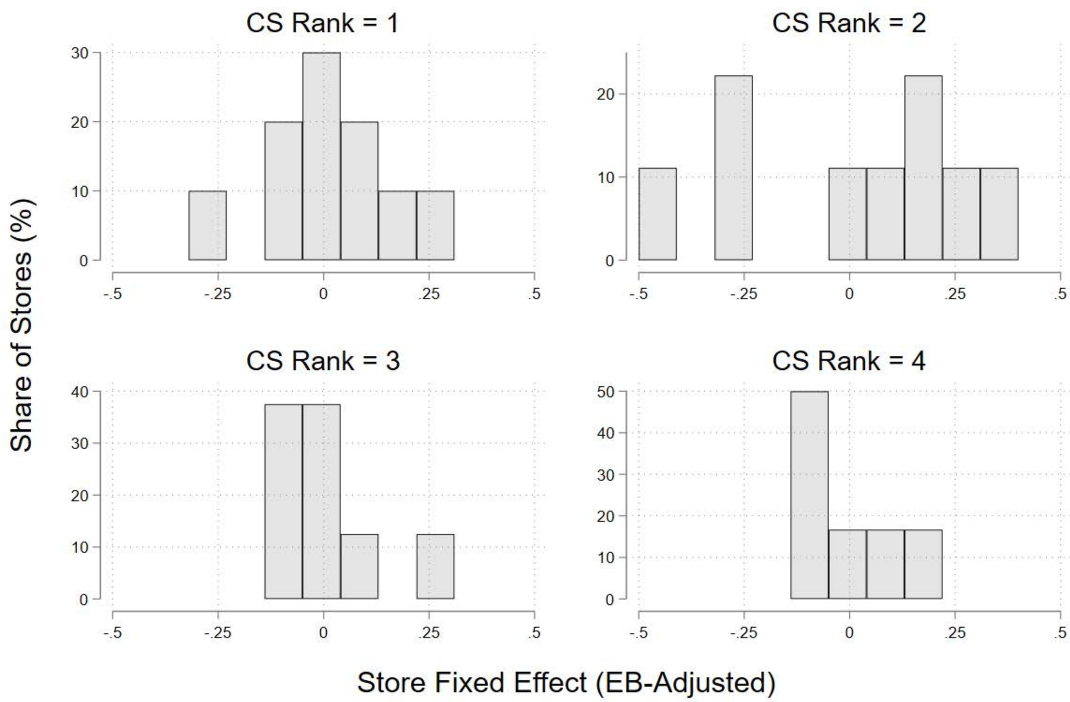
(a) Company A



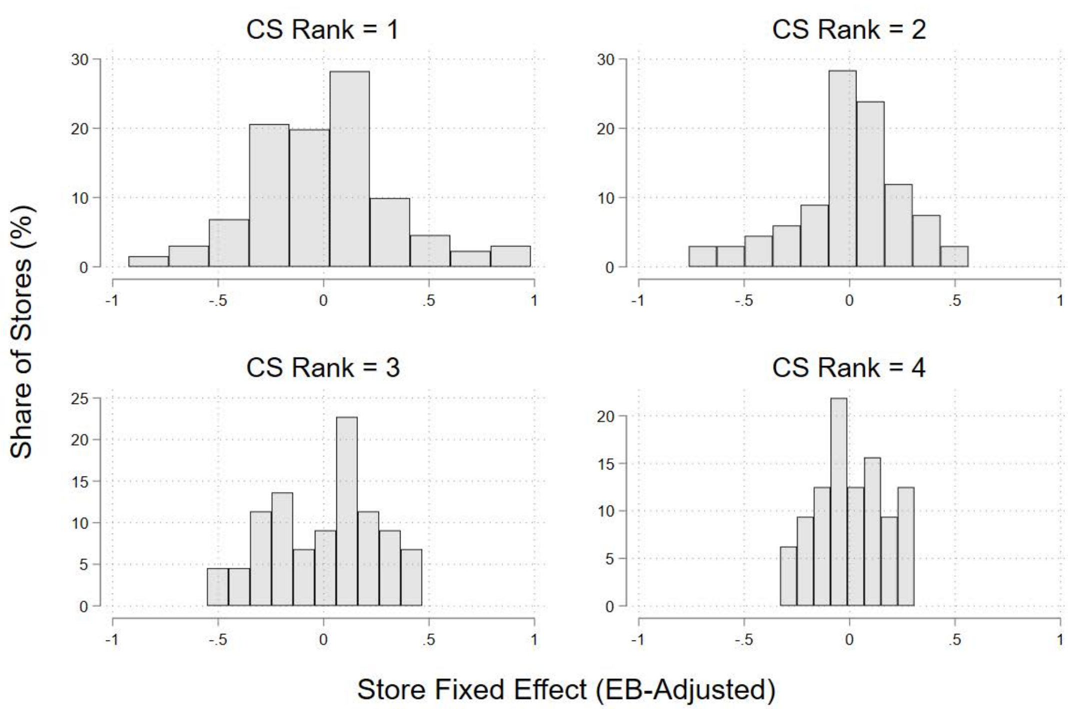
(b) Company B

Note: FE's de-meaned by connected set for plotting.

Figure A.3: Distribution of EB-Adjusted Store Fixed Effects



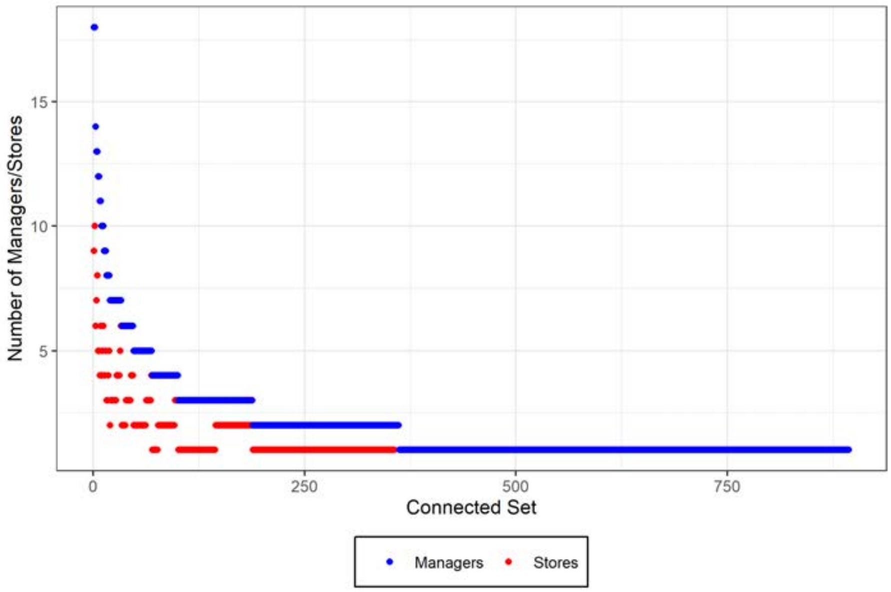
(a) Company A



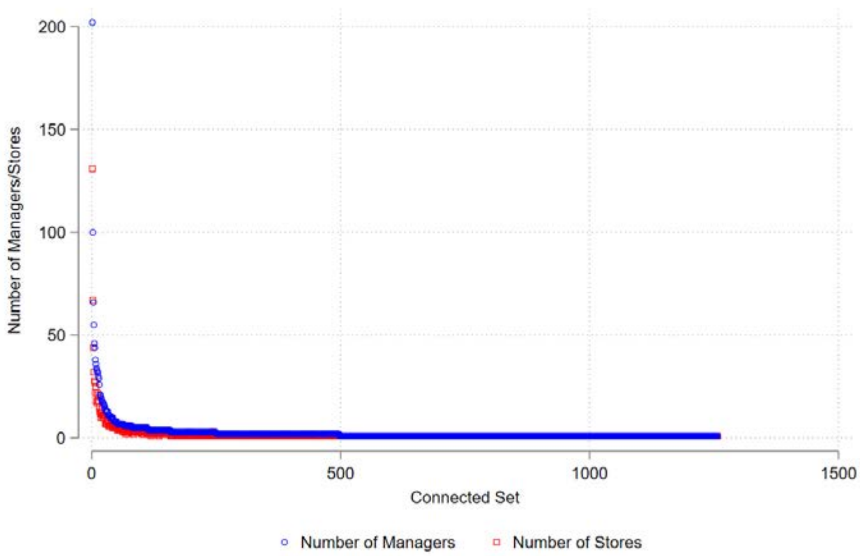
(b) Company B

Note: FE's de-meaned by connected set for plotting.

Figure A.4: Number of Stores and Managers by Connected Set

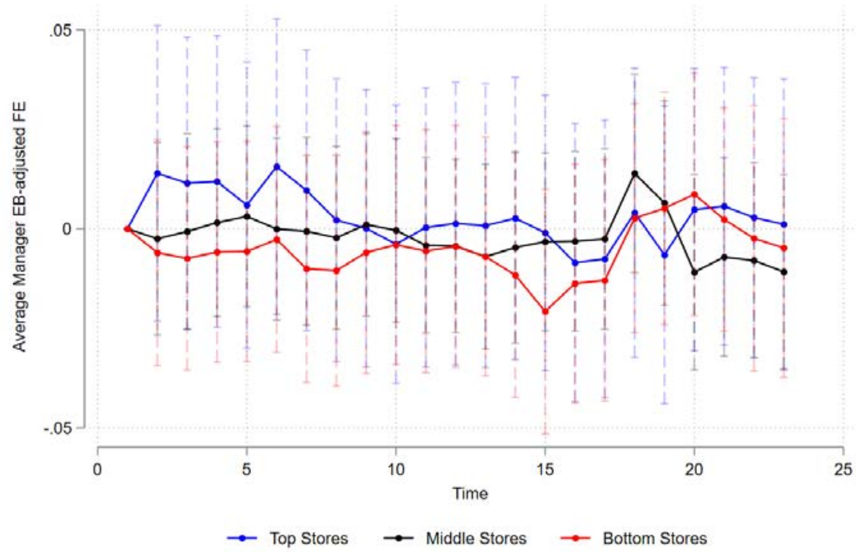


(a) Company A

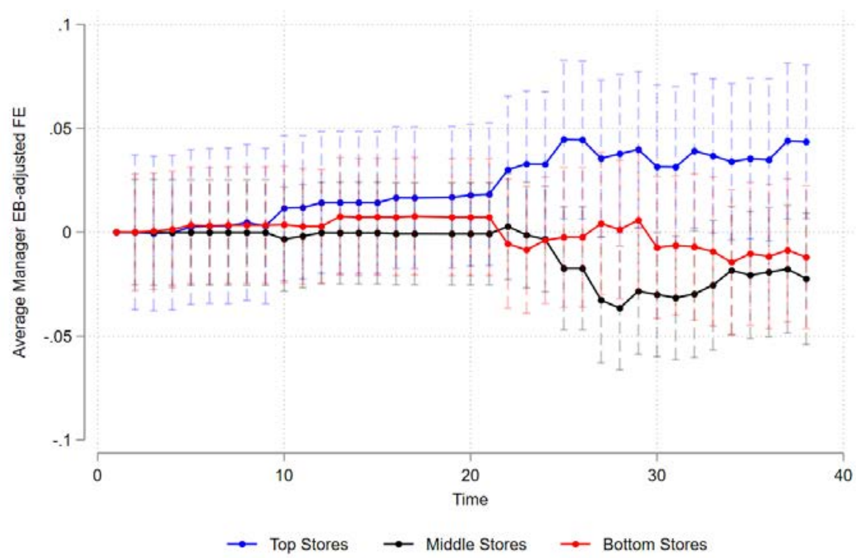


(b) Company B

Figure A.5: Evolution of the Average Manager FE by Store Quality Tercile



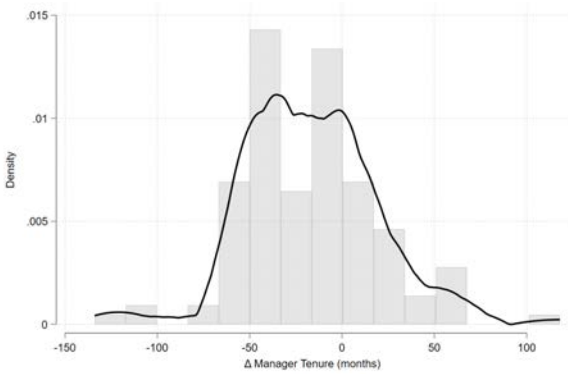
(a) Company A



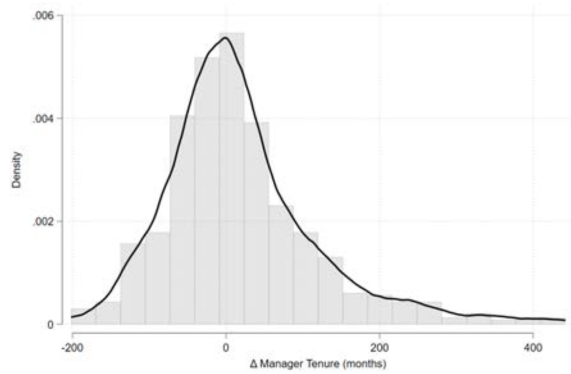
(b) Company B

Note: Charts show the coefficients from regressing the manager quality (EB-adjusted fixed effect) on a time dummy and a connected set fixed effect. Regressions are run separately by store bin (top, middle, bottom), where bins are defined by tercile where each store is located within the store quality (EB-adjusted fixed effect) distribution. Dashed lines represent the 95% confidence interval constructed using robust standard error estimates.

Figure A.6: Distribution of Changes in Manager Tenure

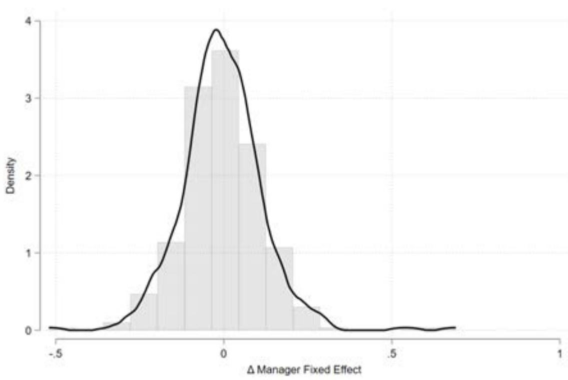


(a) Tenure Change (Company A)

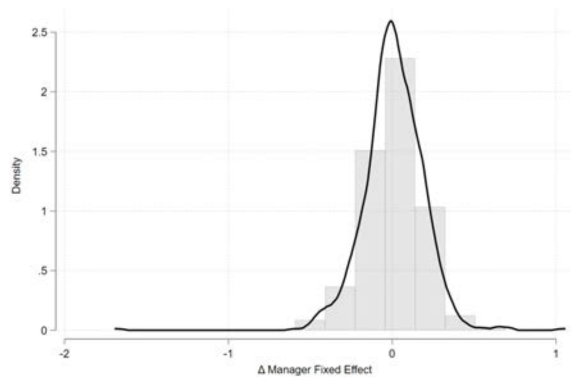


(b) Tenure Change (Company B)

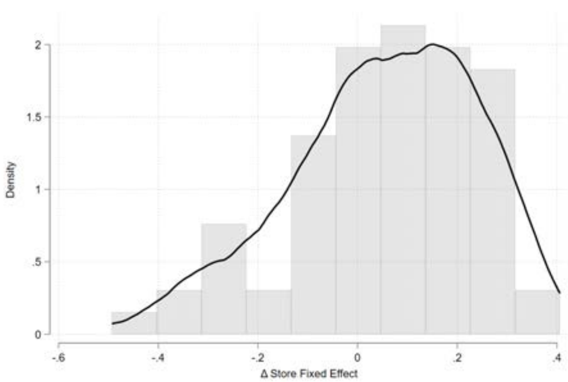
Figure A.7: Distribution of Changes in Manager and Store FE



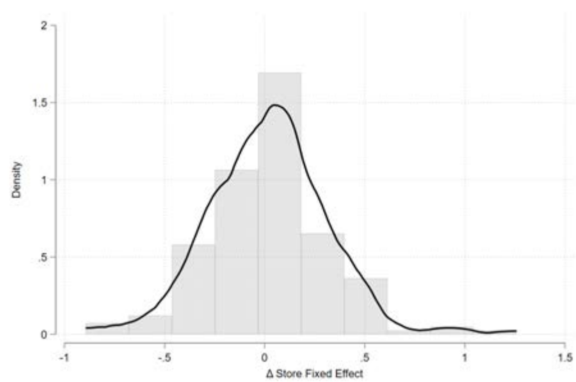
(a) Manager FE (Company A)



(b) Manager FE (Company B)

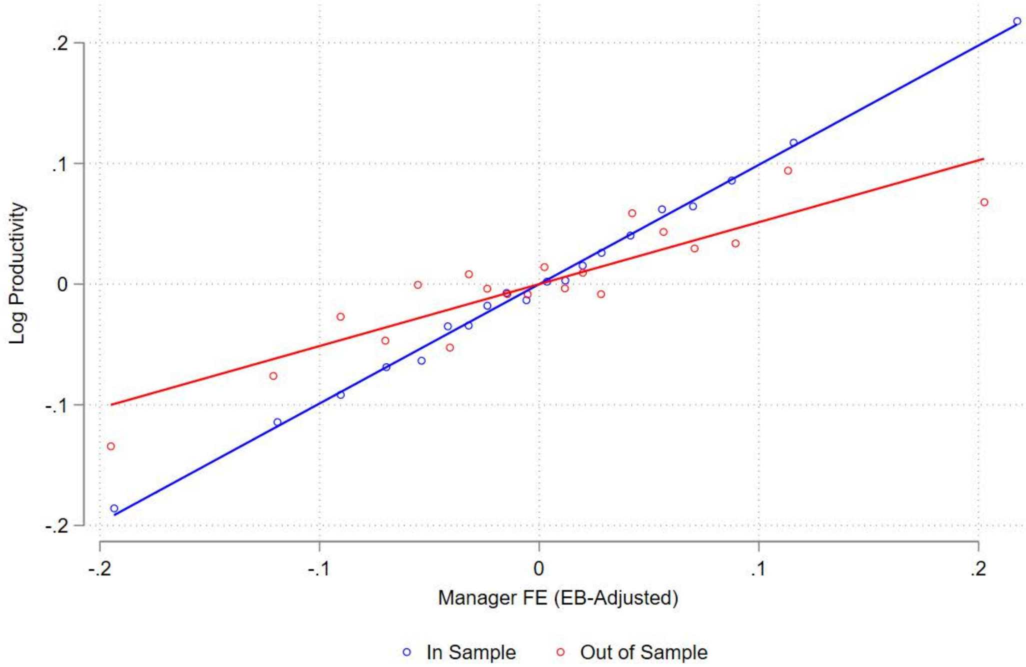


(c) Store FE (Company A)

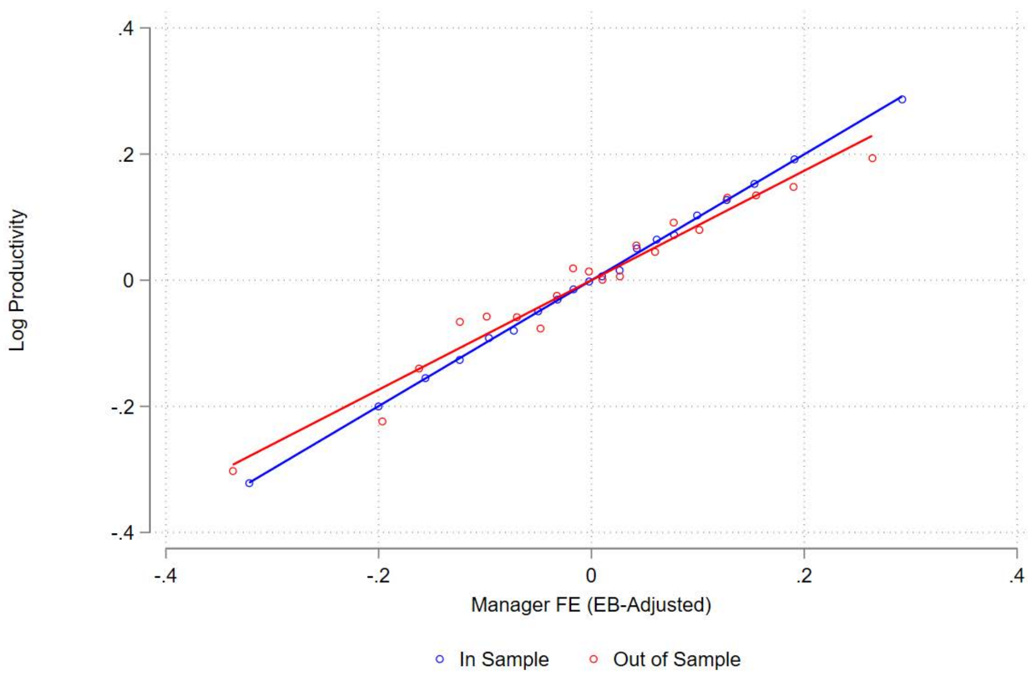


(d) Store FE (Company B)

Figure A.8: Manager FE and Productivity: In and Out of Sample



(a) Company A: In Sample = April 2018 – February 2020 (Before Covid-19);
Out of Sample = March 2020 – December 2020 (During Covid-19)



(b) Company B: In Sample = April 2014 – September 2016;
Out of Sample = October 2016 – May 2017

B Tables

Table B.1: Difference between mover and non-mover stores/managers

	Company A	Company B
% Mover stores	53%	63%
Sales (movers vs. non-movers)	4% lower	0.7% lower
FTE (movers vs. non-movers)	3% lower	6% larger
Store Size (movers vs. non-movers)	1% lower	24% larger
Manager Salary (movers vs. non-movers)	6% lower	
Managers Tenure (movers vs. non-movers)		26% lower

Note: Manager salary is not available for Company B, so we use the manager's tenure at the beginning of the sample instead.

Table B.2: Regression of Female Dummy of Manager and Store Characteristics

	(1)	(2)	(3)	(4)
Mover	-0.067***	-0.054***	-0.031	-0.059
	(0.018)	(0.019)	(0.037)	(0.079)
log Tenure		0.096***	0.110***	0.180***
		(0.012)	(0.027)	(0.048)
Sample	All managers	All managers	Not left	Not left, stores in town > 10
N	3,309	2,955	716	170
Within R^2	0.004	0.035	0.037	0.1147

Note: Robust standard errors are shown in parenthesis. ***, ** and * indicate that coefficients are significantly different from zero at the 1%, 5% and 10% levels, respectively. Columns (2) - (4) control for log(revenue), log(FTE count), log(store area), as well as format (flagship, local, etc) and location (NUTS 2) fixed effects. In the case of movers, those variables are computed for the first store where the manager worked.

Table B.3: Stores Managers Moved To vs Moved From – Company A

Variable	Store	Average	95% CI	
Log Sales	Moved to	12.65	[12.58	12.73]
	Moved from	12.49	[12.41	12.57]
	Difference	0.16	[0.08	0.24]
Log FTE Employment	Moved to	2.56	[2.5	2.63]
	Moved from	2.51	[2.44	2.59]
	Difference	0.05	[-0.01	0.11]
Log Manager Salary	Moved to	8.57	[8.54	8.6]
	Moved from	8.45	[8.40	8.49]
	Difference	0.13	[0.08	0.17]
Log Population Density	Moved to	8.04	[7.81	8.28]
	Moved from	7.89	[7.61	8.17]
	Difference	0.15	[-0.09	0.4]

Note: population density reflects the number of people per square mile (in 2010) in the county where the store is located. Data was obtained from the American Community Survey through IPUMS USA.

Table B.4: List of Observed Shocks and Potential Impacts on Stores

Impact on Stores	**Shocks/Reason for Impact**
Open Late	*Cleaning*
Mid-day Closure (close & reopen in same day)	*Hours Adjustment*
Close Early	Local Authority Evacuation
Full Day Closure	*Local Mandate*
Long-term Closure (> 3 days)	Major Facility Issue
Temporarily Reducing Operating Hrs	Natural Disaster/Weather Condition
Switching to Curbside	Power Outage
Returning to Normal Operating Hrs	Protests
Returning From Curbside	*Staffing Issues*
	Unsafe Conditions
	Other

Note: this table lists all possible impacts on stores and all possible reasons (they are not necessarily matched). Shocks that can also be caused by Covid-19 interventions are shown in *italics*.

Table B.5: Impact of Covid-19 and Manager Experience on Store Sales

	(1)	(2)	(3)	(4)
Manager FE (EB-Adj.)	0.448***	0.447***	0.446***	0.434***
	(0.055)	(0.055)	(0.056)	(0.057)
Store FE (EB-Adj.)	0.997***	0.999***	0.997***	1.003***
	(0.036)	(0.036)	(0.036)	(0.036)
Unsafe Conditions (Covid-19)	-0.383***	-0.502***		
	(0.131)	(0.154)		
Hours Reduction (Covid-19)	-0.195***	-0.011		
	(0.079)	(0.074)		
Manager FE × Unsafe Conditions		0.053***		
		(0.016)		
Manager FE × Hours Reduction		2.209***		
		(0.450)		
Number of Shocks (pre-Covid)			0.017***	0.023***
			(0.005)	(0.006)
Manager FE × Number of Shocks				-0.003**
				(0.001)
N	3,108	3,108	3,108	3,108
Within R^2	0.3230	0.3336	0.2966	0.2977

Robust standard errors in parenthesis and ***, ** and * indicate coefficients are significantly different from zero at the 1%, 5% and 10% levels, respectively. All specifications include a fixed effect for each connected set.

C Estimating Store and Manager Productivity: Challenges and Alternatives

C.1 Alternative Measures of Productivity

There are a number of ways to measure productivity, and in this section we compare our baseline measure with other commonly used alternatives. Our baseline measure sets productivity equal to sales per employee. For the purposes of this section, we call this L-productivity:

$$\text{L-productivity}_{s,t} = \frac{\text{sales}_{s,t}}{\text{FTE employment}_{s,t}}.$$

K-productivity. The first alternative measure is constructed in a similar way, but replacing employment by each store's size. Since the store size is a measure of capital, we refer to this measure as K-productivity:

$$\text{K-productivity}_{s,t} = \frac{\text{sales}_{s,t}}{\text{Floor Area}_{s,t}}.$$

TFPR. The third and final measure we analyze is total factor productivity. Because we do not observe quantities sold, we must construct this measure using revenues and costs, thus TFPR.

We model a store's production function as

$$Y_{s,t} = A_{s,t} L_{s,t}^{\alpha_L} K_{s,t}^{\alpha_K} M_{s,t}^{\alpha_M},$$

where $Y_{s,t}$ is the store's output, $L_{s,t}$ is employment, $K_{s,t}$ measures the store's capital and $M_{s,t}$ measures materials and other inputs to production (overhead costs). Cost minimization implies

$$\frac{\alpha_j}{\sum_j \alpha_j} = \textit{cost share}_j, \quad j \in \{L, K, M\}$$

where *cost share*$_j$ indicates the share of total costs taken by input j (e.g., the labor cost share is *cost share*$_L = \frac{w_t L_{s,t}}{w_t L_{s,t} + r_t K_{s,t} + p_t^M M_{s,t}}$). While the labor and overhead costs are observable, the cost of capital goods for each store must be constructed. We approximate $K_{s,t}$ as the store floor area and r_t as the median rental rate at the county where the store is located.

To simplify computation, we assume constant returns to scale, $\sum_j \alpha_j = 1$.

Given the cost shares α_j, we compute a store's revenue-TFP as[25]

$$TFPR_{s,t} = \frac{(PY)_{s,t}}{L_{s,t}^{\alpha_L} K_{s,t}^{\alpha_K} (p^M M)_{s,t}^{\alpha_M}},$$

where $(PY)_{s,t}$ are sales and $(p^M M)_{s,t}$ are overhead costs of store s in month t.

C.1.1 Relationship Between the Three Measures

Figure C.1 plots the three productivity measures discussed above in a bin-scatter plot (after de-meaning to account for differences in scale), showing that all three of them are very closely related. The same can be seen in Table C.1, which reports a very high correlation between each pairwise combination of the three productivity measures. All of this suggests that our baseline, L-productivity, captures the same features of the store that alternative productivity measures would.

Table C.1: Correlation Between Alternative Measures of Productivity

	L-productivity	K-productivity	TFPR
L-productivity	1		
K-productivity	0.60*	1	
TFPR	0.76*	0.69*	1

Note: * indicates that correlations are statistically significant at the 5% level.

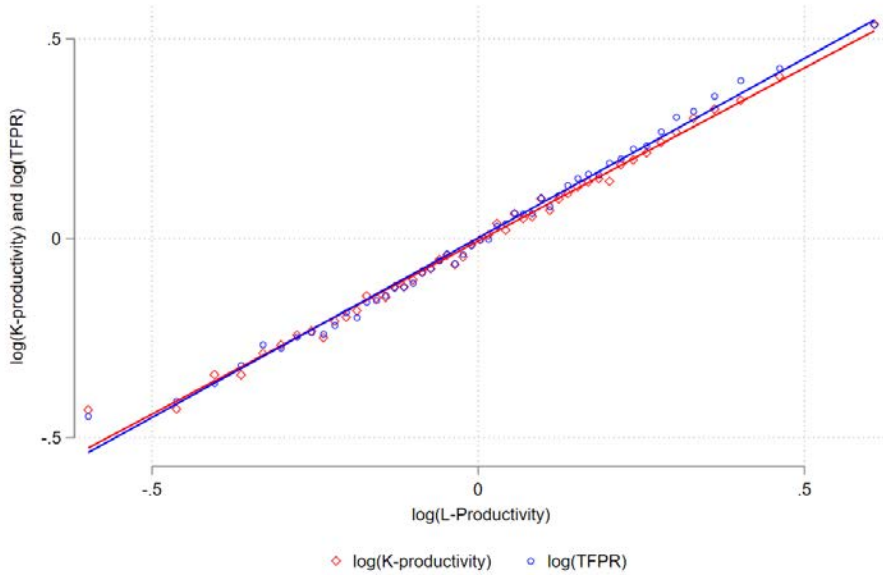
C.2 Allowing for Manager and Store Interactions

Our baseline decomposition of store productivity, equation (1), assumes a log-linear specification that does not explicitly allow for a sorting effect (complementarity) between the stores and managers. To test whether this would change our results, we run the alternative model

$$\log(prod_{s,t}) = \mu_s + \mu_{m(s,t)} + \mu_{s,m(s,t)} + \mu_t + \varepsilon_{s,t}$$

[25] Although we assume a Cobb-Douglas production function, the equation $A_{s,t} = \frac{Y_{s,t}}{L_{s,t}^{\alpha_L} K_{s,t}^{\alpha_K} M_{s,t}^{\alpha_M}}$ holds more generally as a first-order approximation of any production function. In those cases, one can compute α_j as the output elasticity of input j.

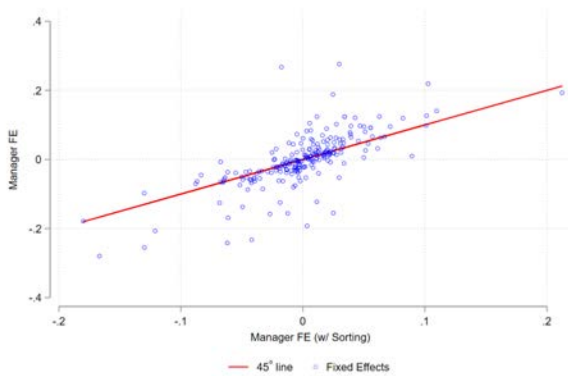
Figure C.1: Relationship Between Alternative Measures of Producivity



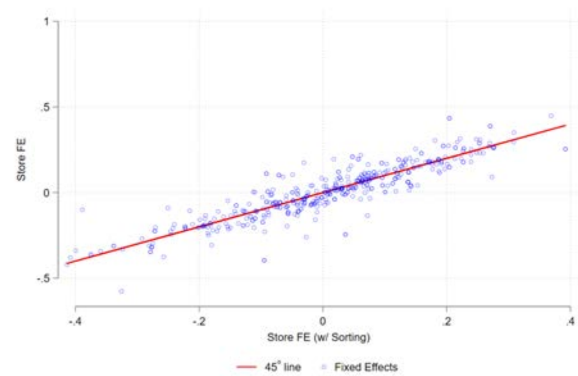
Note: data is binned in to 50 quantiles for plotting; each dot represents the average productivity in its respective bin. All three productivity measures have been de-meaned.

where $\mu_{s,m(s,t)}$ is a store-manager pair fixed effect, which captures any sorting effect between managers/stores. Figure C.2 compares the estimated μ_s and $\mu_{m(s,t)}$ obtained from equation above with the baseline effects estimated in the main text (equation 1). This figure shows that there is little change in the manager and store effects when a term is added to capture any type of complementarity between them.

Figure C.2: Comparing Estimates With and Without Sorting



(a) Manager Fixed Effects



(b) Store Fixed Effects

C.3 Correcting for Measurement Error: Grouped Fixed Effects

A second approach that addresses the inherent measurement error in fixed effect estimation is to abandon the estimation of individual effects and instead focus on groups. In this section, we adopt the two-step procedure proposed by Bonhomme, Lamadon and Manresa (2019). In the first step, we cluster managers and stores into groups with similar units. Then, assuming that all managers and stores in the same group have a common fixed effect, we estimate the influence of each group on an establishment's productivity. The main advantage of this method is that it drastically reduces the number of estimated parameters, raising statistical precision. However, it also inevitably simplifies manager and store heterogeneity and requires us to take a stand on the way units are clustered.

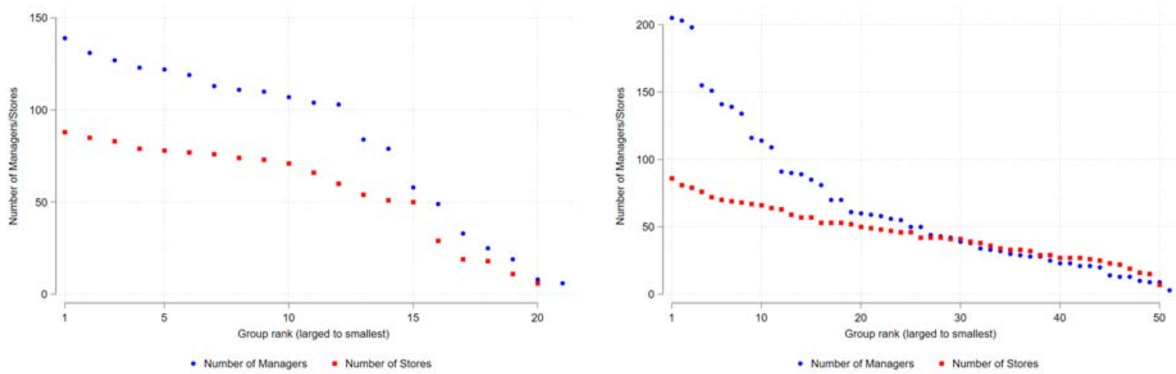
Step 1 We classify both managers and stores into G groups each, using a k-means algorithm. The number of groups is arbitrary, and chosen so that each group contains a relatively large number of managers/stores. We use 20 groups for Company A and 50 groups for company B, so that each groups has, on average, about 50 stores. Varying the number of groups does not meaningfully affect our results.

The key issue that arises in this step is the choice of variables used to cluster the units in our data. Ideally, all of the relevant heterogeneity across managers and stores would be captured by a few observable variables, so that there is little loss of information in the clustering process. We cluster managers in Company A based on their average log-wage across the sample. The intuition is straightforward: all else equal, more productive managers should be paid higher wages. Because this is not available for Company B, we use their average tenure instead.

We cluster stores by their average log FTE employment during the sample, based on the notion that more productive stores should have more employees working in them. While store employment could be guided by company policy, there is scope for managers to at least temporarily change the number of employees if he or she sees fit. This might imply that FTE employment also captures part of the manager's effect, however. To address this concern, we alternatively cluster stores by floor area (a measure of capital in retail) and find very similar results.

Step 2 Once groups are defined, we repeat the estimation of equation 1, with the exception that the effects on the right hand side no longer identify individual managers or stores, but rather the group to which they belong. Figure C.3 shows the number of managers and stores in each group. Figure C.4 below plots the distribution of the estimated grouped fixed effects. In the clustered data, there is a single connected set in the network of stores, so all fixed effects are normalized by the same value.

Figure C.3: Number of Managers and Stores in Each Group



(a) Managers

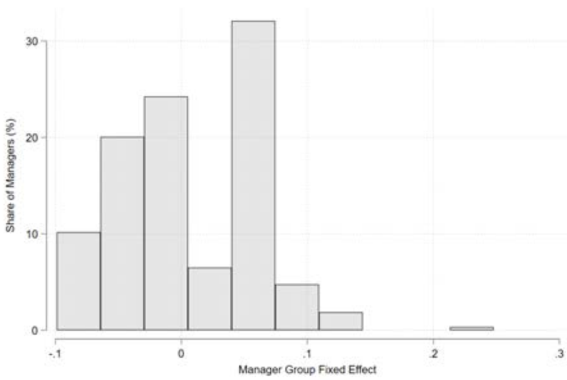
(b) Stores

Comparing Figure C.4 with figures A.2 and A.3, we see that the distribution of grouped fixed effects has most of its mass within a smaller range (this is not obviously the case *a priori*, as the number of connected sets also changes). The shape of the distributions in each of those three figures, however, is relatively similar, even if there are sometimes gaps in the distributions of fixed effects.

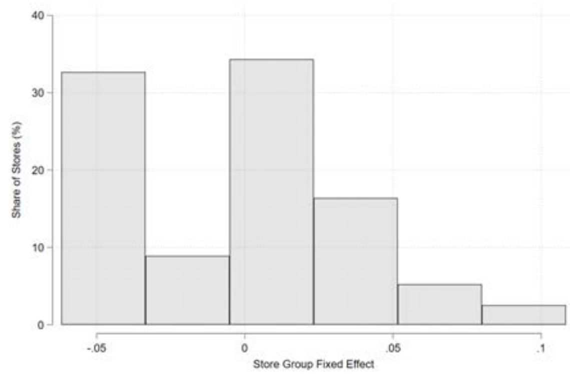
Finally, Figure C.5 plots the relationship between the log of productivity (sales per full-time equivalent employee) and each of the fixed effects we estimate: unadjusted (i.e., the coefficients directly estimated from (1)), EB-adjusted, and grouped. As seen in the figure, the unadjusted and EB-adjusted FEs are very similar, with a noticeable difference only for store quality in Company B. The grouped fixed effects, however, have a significantly flatter relationship with productivity, even if still positive in all cases. This result could indicate that we loose much of the heterogeneity across managers and stores by grouping them along a single dimension.

Figure C.4: Distribution of Grouped Fixed Effects

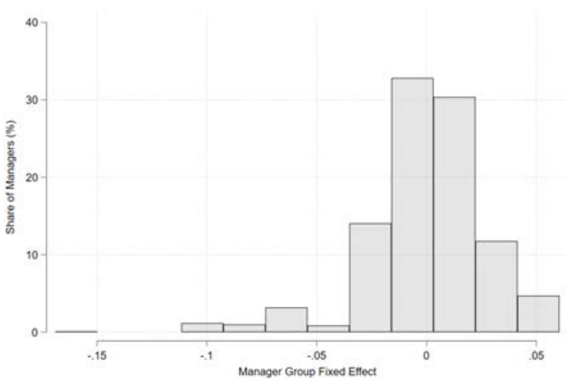
(a) Manager Fixed Effect (Company A)



(b) Store Fixed Effect (Company A)



(c) Manager Fixed Effect (Company B)



(d) Store Fixed Effect (Company B)

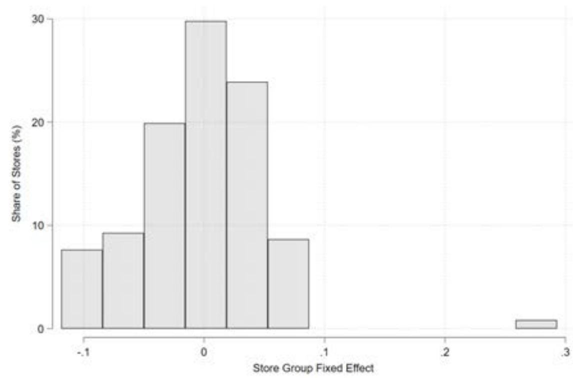
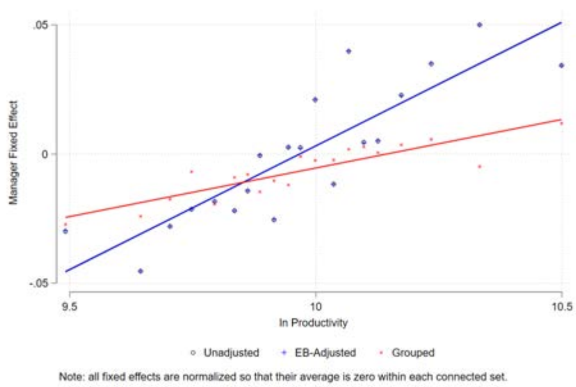
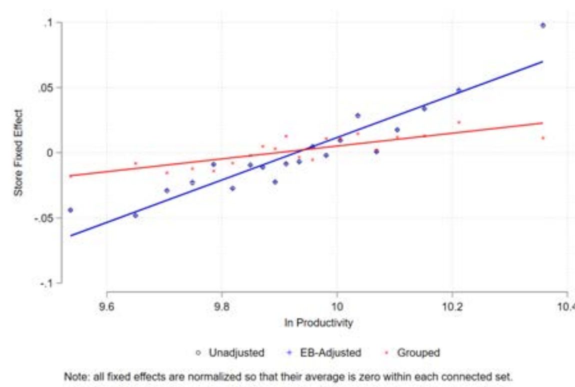


Figure C.5: Unadjusted, EB-Adjusted, and Grouped Fixed Effects

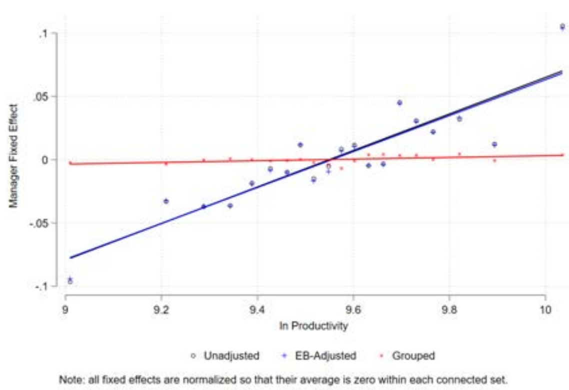
(a) Manager Fixed Effect (Company A)



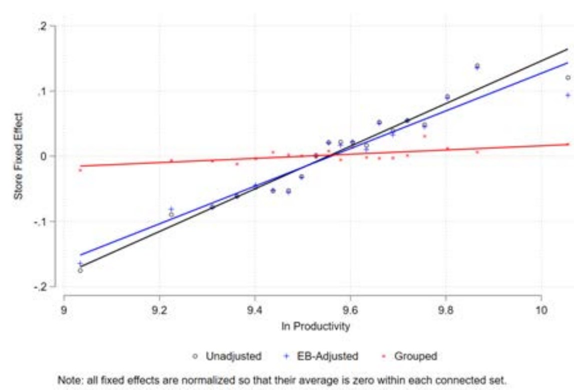
(b) Store Fixed Effect (Company A)



(c) Manager Fixed Effect (Company B)



(d) Store Fixed Effect (Company B)



Note: all fixed effects are normalized so that their average is zero within each connected set.

D Gains from Assortative Matching

As argued in section 3.2.1, there does not appear to be positive assortative matching between managers and stores. We discuss several reasons for why this might be the case, but the fact remains that, under certain assumptions, this is not the optimal allocation of managers to stores. In this section, we compute the potential revenue gain per year should companies assortatively match their managers and stores.

Given model (1), we can do this by simply ranking managers and stores based on their estimated fixed effects, match them based on their ranking, and compute the resulting productivity gain as

$$\Delta \log(prod_{s,t}) = \mu^*_{m(s,t)} - \mu_{m(s,t)},$$

where $\mu_{m(s,t)}$ is the FE of the current manager and $\mu^*_{m(s,t)}$ is the FE of the manager in the optimal match. Given the productivity increase, we can easily calculate the sales revenue increase as well. Because not all managers are present at all times nor can they be compared across connected sets, we only reallocate managers inside the same connected set and period. For Company A, we also differentiate between pre- and post-Covid-19 periods, as the pandemic could have changed the role of managers quite substantially (in which case the pre-Covid estimate is a better reflection of potential gains).

Our estimates are shown in Table D.1, and suggest that Company A could have increased their revenue by 2.22 percent per month in the pre-Covid periods (or the equivalent of about USD 2.23 million/month) simply by optimally allocating their managers *within connected sets* – which means that the unconstrained reallocation could be even more beneficial. During the pandemic, the same exercise results in a slightly lower, but still relevant, estimate: 2.10 percent or USD 1.85 million per month. In Company B the gains are even larger: 6.07 percent gains per month, or about GBP 13.2 million! In each case, our results indicated that are substantial gains from reallocating the best managers to the best stores withing the company.

Table D.1: Sales Revenue Gain per Year with Assortative Matching

	Company A				Company B	
	Pre Covid-19		Post Covid-19			
	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
Percentage	2.22	[1.82, 2.63]	2.10	[1.51, 2.68]	6.07	[5.55, 6.60]
Value (millions)	2.23	[1.90, 2.77]	1.85	[1.33, 2.38]	13.2	[12.0, 14.3]

Note: values are in USD for Company A, and GBP for Company B. Confidence intervals are obtained by bootstrapping sample 100 times.

E Event Study: Pre-Trends and Robustness

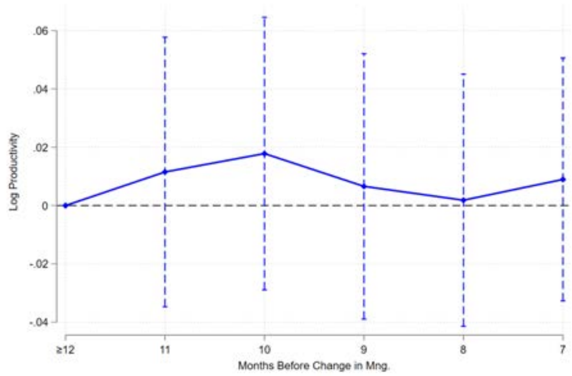
To test for the existence of pre-trends (i.e., before 6 months prior to a change in manager) in out study, Figure E.1 runs model (3') in the periods used to match treated to control stores. As expected, none of the coefficients are statistically different from zero at the usual confidence levels, suggesting that stores were in a common trend before the change in manager became apparent.

E.1 Managers who Stay, and Managers who Leave

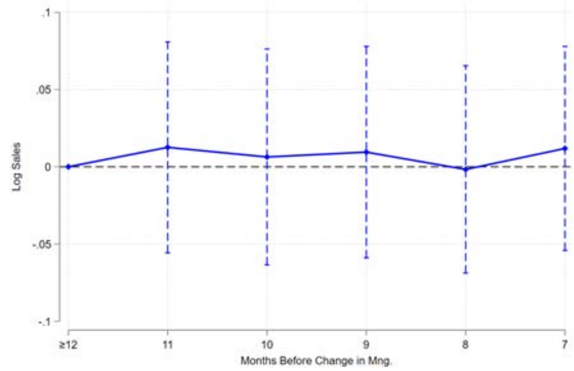
As mentioned in section 4, there are many reasons why a manager could leave a store. Among those, managers who leave a store but remain within the company (that is, move to another store) might be very different from those that leave the Company altogether (are fired, move to a different job). We remove managers that leave the company after a tenure longer than 20 years to account for retirees. Figure E.3 tests this hypothesis by conducting each of the event studies on the subset stores whose initial manager stays in the company.

Most of the effects we have seen in figures 5 and 6 are also present in for managers who switch stores within the company, but are generally greatly attenuated. This is particularly true for Company A (Figure E.3), where are all estimated coefficients are considerably closer to zero than in the full study – note in particular that the drop in wages is almost half as big as the drop in Figure 5 (firms would plausibly not want to punish managers that transferring internally). The results for Company B (Figure E.4) show that stores with managers that move within the company experience a less pronounced decline in productivity and sales before the change in manager. We also see a higher

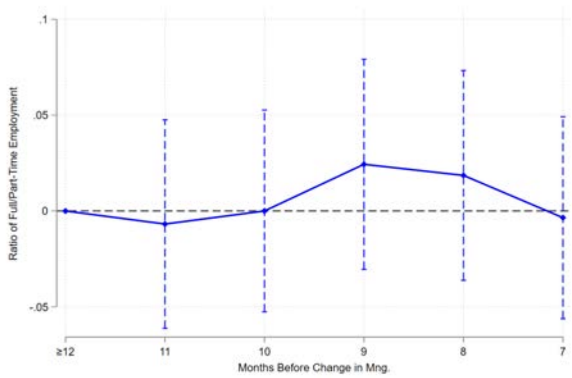
Figure E.1: Event Study: Outcomes During Matching Periods
Company A



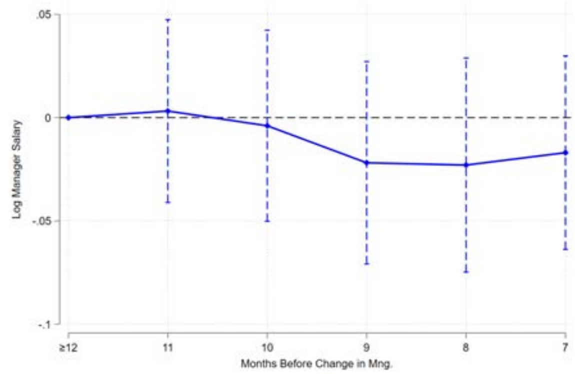
(a) Productivity



(b) Sales



(c) Full/Part-Time Emp. Ratio



(d) Manager Salary

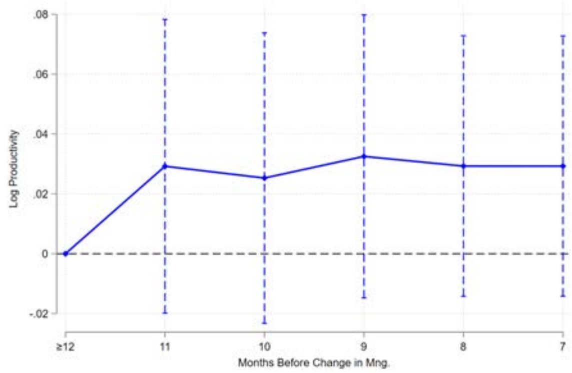
Note: dots show the value of each coefficient β_k, while whiskers indicate the 95% confidence interval (computed using robust standard errors).

increase in the ratio of full-time to part-time workers after the managerial change, although coefficients are not statistically significant (we should note that Company B has much fewer managers that move internally, relative to the full set of movers).

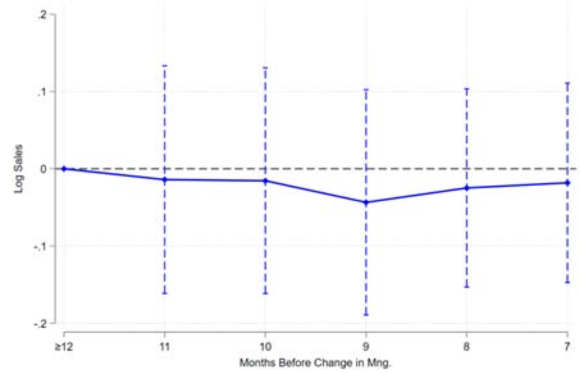
E.2 Switching From Bad to Good Managers

Section 4 shows that the average difference in quality of the incoming and outgoing manager across all stores is close to zero. In this section, we focus on what happens to stores when the difference is big: specifically, when stores move from a bad to a good manager, or vice-versa. We classify managers into top and bottom based on their position relative to median of their connected set. Figures E.5 Figures E.6 then repeat the event study

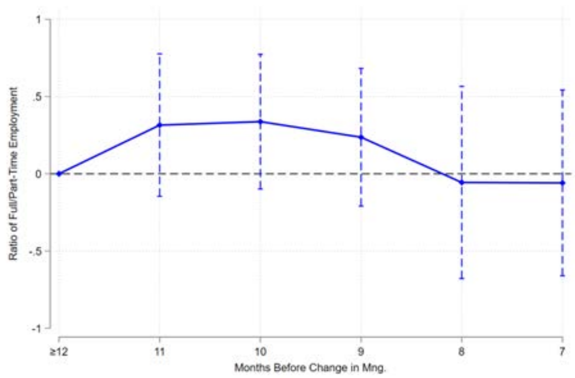
Figure E.2: Event Study: Outcomes During Matching Periods
Company B



(a) Productivity



(b) Sales

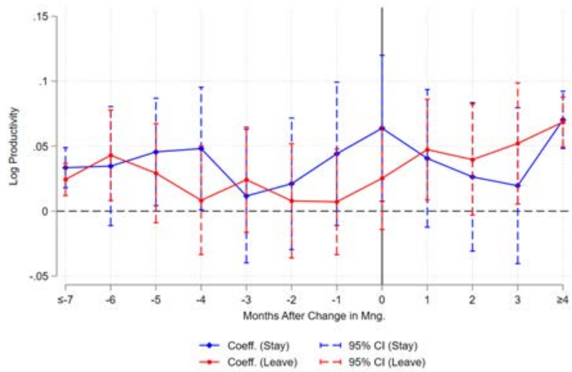


(c) Full/Part-Time Emp. Ratio

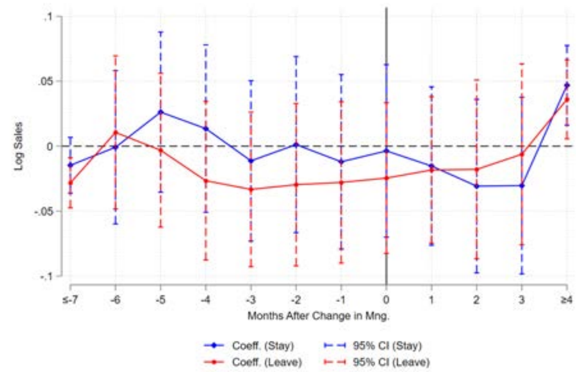
Note: dots show the value of each coefficient β_k, while whiskers indicate the 95% confidence interval (computed using robust standard errors).

analysis for each group of stores (top-to-bottom, and bottom-to-top). It is quite evident in both figures that after a store moves from a bottom to a top manager, it experiences an increase in productivity, while the opposite is true when it moves from a top to a bottom manager. In company A, we can also see that the impact on the wage of top managers right before they leave is less steep than the impact on the wages of bottom managers. In terms of sales and employment, top also managers do better than bottom managers: the blue line trends above the red line before time 0, with a reversal or catch up in the latter periods when stores switch managers.

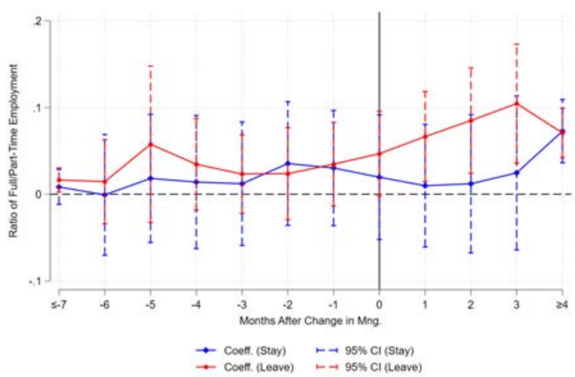
Figure E.3: Event Study: Managers that Stay vs Managers that Leave (Company A)



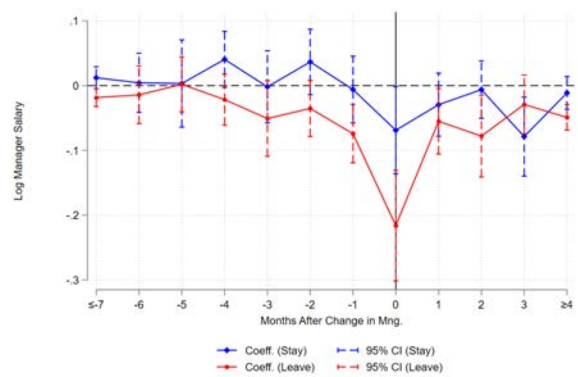
(a) Productivity



(b) Sales



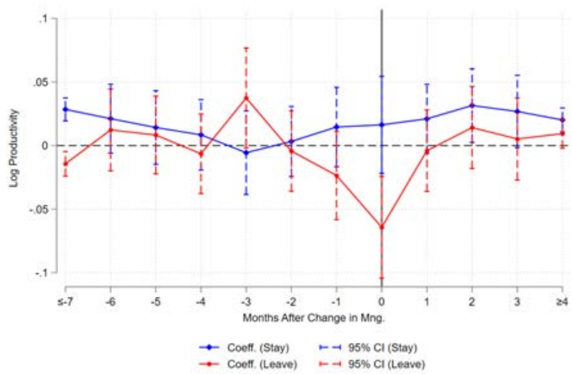
(c) Full/Part-Time Emp. Ratio



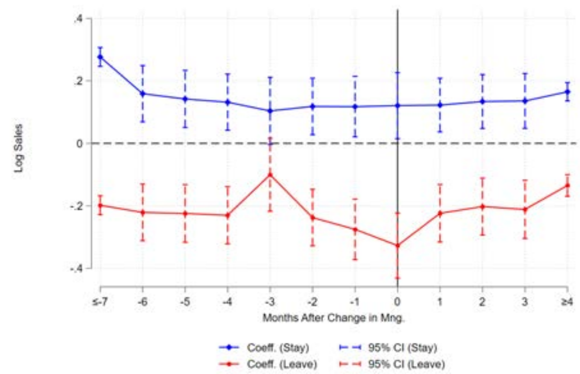
(d) Manager Salary

Note: dots show the value of each coefficient β_k, while whiskers indicate the 95% confidence interval (computed using robust standard errors). The red line represents coefficients for managers that leave the company after the event; the blue line represents managers that move store but remain within the company.

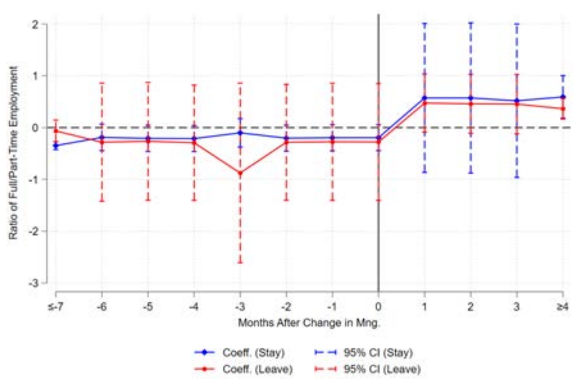
Figure E.4: Event Study: Managers that Stay vs Managers that Leave (Company B)



(a) Productivity



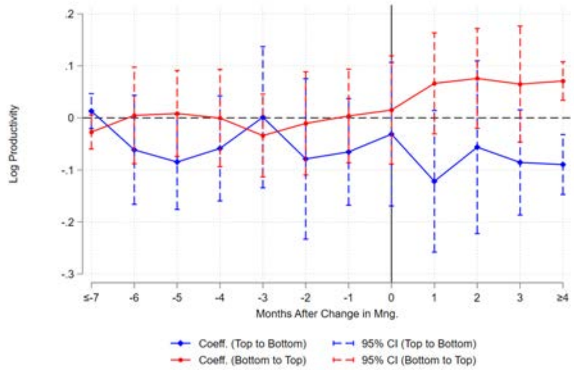
(b) Sales



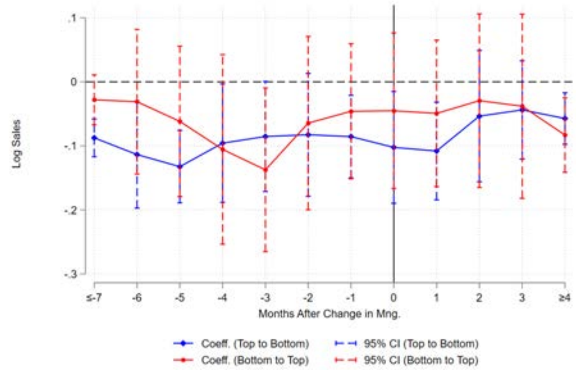
(c) Full/Part-Time Emp. Ratio

Note: dots show the value of each coefficient β_k, while whiskers indicate the 95% confidence interval (computed using robust standard errors). The red line represents coefficients for managers that leave the company after the event; the blue line represents managers that move store but remain within the company.

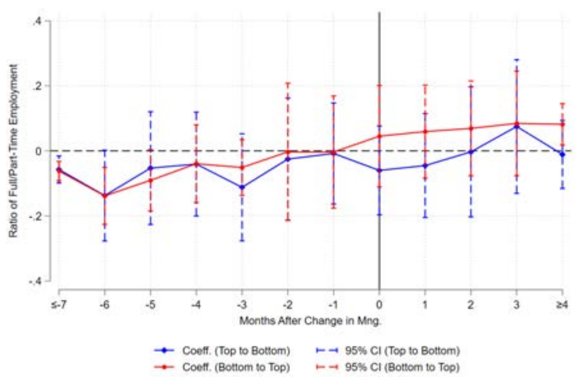
Figure E.5: Event Study: Switching Bottom and Top Managers (Company A)



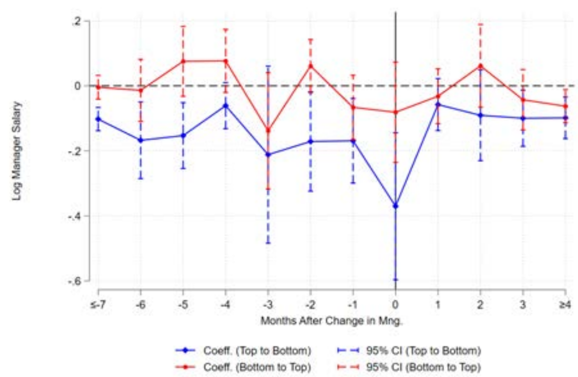
(a) Productivity



(b) Sales



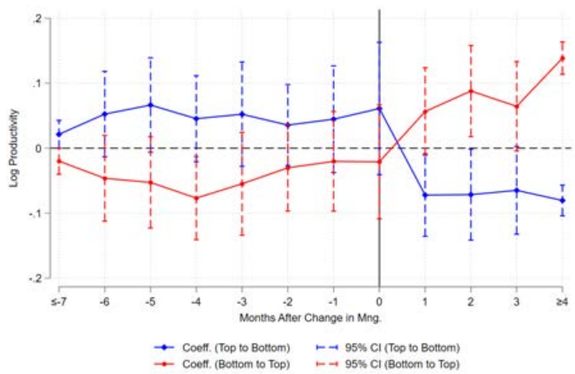
(c) Full/Part-Time Emp. Ratio



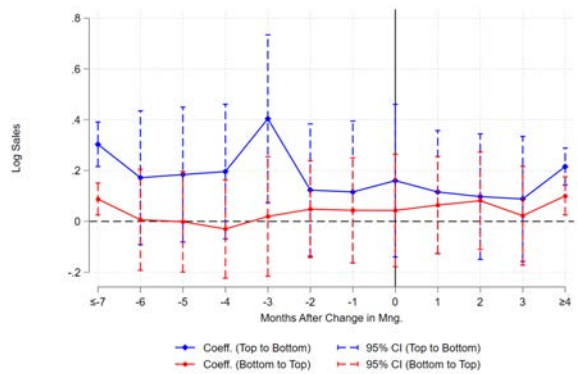
(d) Manager Salary

Note: dots show the value of each coefficient β_k, while whiskers indicate the 95% confidence interval (computed using robust standard errors). The red line represents coefficients for stores that switch from bottom (below median) to top (above median) managers; the blue line represents stores that switch from top to bottom managers.

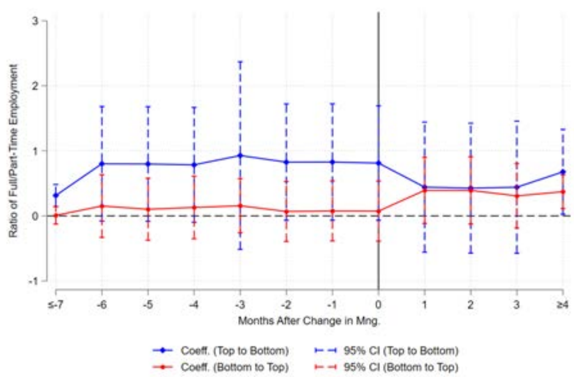
Figure E.6: Event Study: Switching Bottom and Top Managers (Company B)



(a) Productivity



(b) Sales



(c) Full/Part-Time Emp. Ratio

Note: dots show the value of each coefficient β_k, while whiskers indicate the 95% confidence interval (computed using robust standard errors). The red line represents coefficients for stores that switch from bottom (below median) to top (above median) managers; the blue line represents stores that switch from top to bottom managers.

F Manager Productivity and Energy Usage

In this section, we explore the impact of managers on their store's energy consumption. We do that by regressing energy usage on the manager's productivity (EB-adjusted fixed effect), as well as other store-level controls, and exploring the signs of the correlations we find. We focus on electricity consumption in stores in Company B (we do not have energy use data for company A), as electricity use is almost universal across stores. In addition to that, about 40% of stores in our sample also use gas as a source of energy. Those stores tend to be concentrated in specific regions of the UK (perhaps reflecting local differences in the energy grid), are larger (in terms of store area), but also have lower revenues than similar stores that use solely electricity. Importantly, we find no correlation between the manager EB-adjusted fixed effect and the likelihood that a store uses gas. In any case, because electricity and gas might be substitutes, our regression includes a store-level dummy variable that equals 1 if a store records positive gas usage in at least one period of our sample.

We present the results in Table F.1. Note that all specifications include time and location (city) fixed effect to capture seasonal and geographical movements in electricity consumption, connected set fixed effects to absorb differences in average manager and store productivity across connected sets, as well as store-level dummy variable that indicates each period in which a store adopts energy saving technologies, such as chillers and LED lights (energy tech FE).[26] We find that a store's electricity consumption increases with its size, both in terms of floor area and revenue, and decreases as the price of electricity increases.[27] Those findings are intuitive, although we also highlight that there are clear endogeneity concerns in the coefficients, particularly on electricity prices.

We also find that more productive managers tend to consume less electricity, but that there is no relationship between electricity consumption and store productivity. Those results hold regardless of movements in energy prices, or the size of the store. Columns (2) and (3) show that they remain robust to the availability of alternative energy sources (gas), which has a negative impact on electricity consumption as gas substitutes for elec-

[26] We crudely estimate the effect of energy savings technologies on energy usage by regressing log electricity consumption on the energy tech dummy, while controlling for electricity prices, average temperature, as well as time and store fixed effects. We find that adopting such technologies brings electricity consumption down by about 1.3 percent each month. This is consistent with results from Kahn and Kok (2014), who argue that newer buildings tend to be more energy-efficient than older ones.

[27] Energy prices reflect the average price of energy consumption, calculated as energy cost/energy consumption (both per month).

tricity; the number of hours a each store is open per week, which has a large and positive effect on energy consumption; and the average monthly temperature in the store's vicinity. Average monthly temperature has a U-shaped impact on energy use, indicating that very low or very high temperature are associated with higher energy use (presumably for heating and cooling). However, column (4) shows that more productive managers can attenuate this effect, with a smaller increase in energy consumption during the colder and warmer months (on top of an already negative baseline impact on energy use).[28]

[28]Average monthly temperature is obtained from measurements by the weather station that is closest to each store. We obtain data from each station from the Met Office website. Using the geographical coordinates for both stations and stores, we calculate the geodesic distance between each station-store pair and match each store to its nearest station. The average distance between a store and its nearest station is about 0.4 miles, with all stores located within 2 mile distance from its nearest station

Table F.1: Impact of Manager Productivity and Store Characteristics on Electricity Usage

	(1)	(2)	(3)	
Manager FE	-0.169***	-0.115**	-0.114*	-0.147**
	(0.063)	(0.058)	(0.058)	(0.569)
Store FE	-0.130	-0.079	-0.079	- 0.079
	(0.095)	(0.094)	(0.094)	(0.094)
Log Electricity Price	-0.695***	-0.775***	-0.778***	-0.775***
	(0.161)	(0.149)	(0.148)	(0.148)
Log Revenue	0.500***	0.356***	-0.355***	0.356***
	(0.030)	(0.029)	(0.029)	(0.029)
Log Store Area	0.578***	0.640***	0.640***	0.640***
	(0.022)	(0.023)	(0.023)	(0.023)
Uses Gas		-0.103***	-0.103***	-0.103***
		(0.028)	(0.028)	(0.028)
Log Hours Open (week)		0.965***	0.967***	0.966***
		(0.118)	(0.118)	(0.118)
Log Avg. Temperature			-0.527***	-0.608***
			(0.130)	(0.147)
(Log Avg. Temperature)^2			0.118***	0.137***
			(0.041)	(0.044)
Log Avg. Temperature × Manager FE				0.037**
				(0.017)
(Log Avg. Temperature)^2 × Manager FE				-0.010**
				(0.004)
N	19,252	19,162	19,162	19,162
Within R^2	0.8155	0.8324	0.8327	0.8329

Manager and Store FE are EB-adjusted. Standard errors are clustered at the connected set level and shown in parenthesis. ***, **, and * indicate coefficients are significantly different from zero at the 1%, 5% and 10% levels, respectively.

G Agglomeration Economies and Store Quality

There is a well-known relationship between population density and firm-level productivity (Duranton and Puga, 2004; Combes et al., 2012; Gennaioli et al., 2013; Behrens, Duranton and Robert-Nicoud, 2014). As mentioned above, one reason why this relationship is not seen in Table 3 is that the connected set fixed effects absorb most of the variation in productivity across locations. For that reason, we also estimate an alternative regression where store level productivity is directly regressed on log population density and log income per capita. Because both of these variables vary little over time, we first average all required variables in our data over time to focus on cross-sectional variation at the store level. As before, observations are weighted by the number of periods they are present in the data. For Company B, a dummy for store format is also included.

The results in Table G.1 present novel evidence that agglomeration effects can be observed *within firms*, not just across firms or across cities. Column (1) shows a positive and statistically significant relationship between store-level productivity and both log population density and income per capita. This relationship becomes much weaker (in Company B) or less precise (Company A) for population density in column (2), when we include connected set fixed effects and restrict the sample to the stores for which we separately identify store and manager quality (making it comparable to results in tables 2 and 3). This result is in line with the observation that connected sets tend to contain stores that are relatively more similar, including being located in cities of similar size.

Finally, Column (3) includes the store quality itself as a regressor. Again there is some heterogeneity across companies, but in both cases we find that the coefficient on population density is not statistically significant and that the coefficient on income per capita drops considerably. This indicates that any residual effect of population density on productivity (e.g., better stores are in larger cities) is captured by the store quality. In contrast, log income per capital remains positive and significant throughout, suggesting that it can impact productivity through other channels (for example by increasing sales due to higher consumer demand).

It is important to note that the same patterns hold true if we first collapse the data to the *manager* level (not the store level, since managers move across stores) and then include the manager quality as a control. As a result, the results shown here indicate that agglomeration economies (including sorting, selection, and other patterns related to population density) are (1) positively correlated with variation in productivity even within companies; and (2) that those effects are mostly captured by the store or manager quality measures we estimate from equation (1). However, there are still other channels through which local variables can affect productivity that are not completely absorbed by manager and store quality – specifically, features that increase local average income without necessarily increasing population density.

Table G.1: Regression of EB-Adjusted Store Fixed Effects on Observable Characteristics

	Company A			Company B		
	(1)	(2)	(3)	(1)	(2)	(3)
log population density	0.018**	0.020*	0.004	0.008**	0.002	0.007
	(0.007)	(0.012)	(0.010)	(0.004)	(0.006)	(0.006)
log income per capita	0.089**	0.311***	0.144***	0.188***	0.167***	0.125***
	(0.037)	(0.072)	(0.057)	(0.024)	(0.046)	(0.057)
EB-adjusted store FE			0.643***			0.467***
			(0.084)			(0.044)
Restricted Sample	No	Yes	Yes	No	Yes	Yes
Connected Set FE	No	Yes	Yes	No	Yes	Yes
N	857	278	278	1,989	906	906
Adjusted-R^2	0.022	0.159	0.681	0.078	0.106	0.497

'Restricted Sample' indicates that the sample only includes stores for which both store and manager quality are separately observable; regressions for Company B also include store format (flagship, local store, etc.) dummies. Standard errors are clustered at the location level and shown in parenthesis. ***, ** and * indicate that coefficients are significantly different from zero at the 1%, 5% and 10% levels, respectively.