Sources of Firm Life-Cycle Dynamics: Size vs. Age Effects*

Lorenz Kueng†

Mu-Jeung Yang‡

Bryan Hong§

January 2017

Abstract

What determines the life-cycle of businesses? Exploiting unique firm-level panel data on internal organization and innovation we establish three key sets of stylized facts to inform recent theories of firm life-cycles. First, life-cycle effects are driven by startups, not by new establishments of existing firms. Second, organizational restructuring and innovation are both strongly correlated with firm growth but not with firm age, in contrast to passive learning theories of firm dynamics. Third, there are important sectoral differences in innovation activities which are monotonically increasing in firm size for manufacturing firms but hump-shaped for firms in service industries.

*The usual disclaimer applies. An earlier version of results of this paper was circulated under the title "The Age Dynamics of Productivity, Management Practices and Organizational Capital." We would like to thank Nick Bloom, Chad Jones, Pete Klenow, Josh Lerner and seminar participants at Stanford and Harvard Business School for useful comments.

†Universita della Svizzera italiana (USI) , and NBER

‡Corresponding Author: mjyang@eccles.utah.edu, David Eccles School of Business, University of Utah

§Bloch School of Management, University of Missouri

1 Introduction

A large empirical literature has documented important age effects in the growth and exit dynamics of businesses.[1] Moreover, recent empirical evidence suggests that establishments follow distinct life-cycle patterns which are crucial for understanding aggregate outcomes. For instance, Haltiwanger, Jarmin and Miranda (2013) show that establishment age dynamics account for a major part of employment growth and Hsieh and Klenow (2014) argue that these firm dynamics are also important for understanding cross-country differences in aggregate productivity. Not surprisingly, a large theoretical literature has proposed a variety of mechanisms to explain life-cycle dynamics of businesses, with particular emphasis on organizational capital accumulation and innovation. Although these theories are successful in capturing business growth patterns over the life-cycle, direct empirical evidence on the mechanisms described by these theories remains scarce.

We take advantage of unique firm- and establishment-level panel data on internal organization and innovation, representative of the entire Canadian economy (including non-manufacturing firms) to study recent theories of firm life-cycle patterns. This paper makes two main contributions. First, we show that most of the life-cycle dynamics observed in the data are driven by startups rather than new establishments of existing firms. In other words, the objects of interest are firm life-cycles rather than establishment life-cycles, so that we focus our attention to startups. Second, we explore various mechanisms driving startup life-cycle dynamics with a particular emphasis on organizational restructuring and innovation. To inform theories of firm life-cycle patterns we exploit the fact that models of endogenous productivity can be grouped into those that emphasize firm size effects and those that emphasize firm age effects as drivers of firm life-cycle dynamics. Size

[1] See for example the empirical facts provided by Dunne, Roberts and Samuelson (1989) and the survey by Sutton (1997).

effects are defined as systematic changes in internal organization or innovation as firms grow. We argue that such size effects naturally emerge in a variety of models of endogenous organizations that have been popular to analyze firm dynamics, such as Calvo and Wellisz (1978), Aghion and Tirole (1997), Garicano (2000), Caliendo and Rossi-Hansberg (2012) or Akcigit, Alp and Peters (2014). In contrast, age effects are defined as systematic changes in internal organization or innovation as firms get older. We show that age effects are closely related to models of "passive learning" in which firms need time to learn about parameters of their economic environment, as for example in Jovanovic (1982). Related models that generate conditional age effects are Rajan and Zingales (2001),Acemoglu, Aghion, Lelarge, Van Reenen and Zilibotti (2007) and Garicano and Rossi-Hansberg (2012).

Whether life-cycle dynamics of organizational restructuring and innovation are driven by size or age effects matters for a least two distinct reasons. First, an empirical exploration of firm age and size effects is informative about the deeper determinants of startup life-cycle growth and the explanatory power of passive learning models. Second, this deeper understanding of startup growth directly matters for the effectiveness of external financing and expertise, frequently provided by financial intermediaries such as venture capital. If life-cycle dynamics are mainly related to size effects, then external financing and expertise on how to swiftly scale up organizations can be powerful accelerators of startup growth as suggested by Hellmann and Puri (2002) and Puri and Zarutskie (2012). On the other hand, if age effects are crucial, then there are limits to speeding up startup growth, despite the availability of external financing and expertise.

Our analysis is made possible by a unique dataset on internal firm organization and innovation that is representative of the entire Canadian economy and spans the time between 1999 and 2006. The data include detailed information on various organizational dimensions and management practices that are at the heart of recent theories of organizational capital accumulation, such as organizational layers, centralization, span of control, and performance pay. Additionally, our dataset provides detailed measures on innovation activities which allows us to capture product, process and incremental innovations, thereby providing a richer picture of innovation activity than just R&D or patenting, which are mainly pursued by very large and mature firms. Importantly, the data is rich enough to let us focus on startups and contrast age and size effects of manufacturing firms with service firms. Methodologically, we include a full set of firm fixed effects, so that all variation in our estimated age and size effects is estimated off within-firm variations of organizational restructuring and innovation.

We highlight three key sets of results on the life-cycle dynamics of startups. First, among a variety of measures for management practices and innovation, we fail to find evidence for conditional age effects. In other words, conditional on firm size, firm age does not seem to have any direct influence on organizational restructuring or innovation.

Second, a number of novel patterns for size effects emerge that are generalizable across sectors. Among these are important non-monotonic size effects for organizational layers. We find that the number of layers tends to increase in firm size, consistent with Caliendo, Monte and Rossi-Hansberg (2015). However, we also show that for firms growing past 300 employees there is a systematic tendency of delayering. In other words, the number of organizational layers follows a stark inverted U-shape pattern that might provide an interesting starting point for new models of firm dynamics with endogenous layers.

Third, we find a number of important differences in size effects when comparing organization restructuring and innovation of manufacturing and service firms. For example, while service firms exhibit a systematic increase in centralization of decision making and increasingly adopt performance pay as they grow bigger, we find no such relationship among manufacturing firms. Importantly, we document stark differences in innovation size effects between manufacturing and service firms. Manufacturing firms that grow bigger are also more likely to generate product, process and incremental innovations, consistent with endogenous innovation models such as Klette and Kortum (2004). In contrast, service firms exhibit a non-monotonic relation between innovative activity and firm size. As service firms grow bigger, they first become more innovative before turning significantly less innovative as they grow very large. These correlations are indicative of very different types of scale effects in innovation production functions applying to service industries compared to manufacturing.

This paper is closely related to empirical work on the drivers of startup and life-cycle growth patterns. Much of the attention in this literature has either focused on the role of external financing as in Evans and Jovanovic (1989), Holtz-Eakin, Joulfaian and Rosen (1994), Hurst and Lusardi (2004) and Black and Strahan (2002) or the role of venture capital in promoting startup growth as in Hellmann and Puri (2002) and Puri and Zarutskie (2012). In contrast, we focus on how organizational restructuring and innovation correlates with firm age and firm size, and which economic mechanisms potentially drive startup life-cycles. Our work is therefore complementary to Ouimet and Zarutskie (2014) who analyze the workforce age composition of startups over the firm life-cycle.

2 Theories of Firm Life-Cycle Dynamics

As background for our empirical analysis, we start by providing an overview of recent theories of firm age dynamics. We focus on theories of the actual mechanisms behind the accumulation of organizational capital such as organizational restructuring and innovation. This review is not intended to be comprehensive but will rather reflect availability of direct measures of these theories in our data.

One can roughly classify age dynamics theories into two groups following two popular approaches to firm dynamics. First, following models of optimal organization and "active learning" as in Ericson and Pakes (1995), Klette and Kortum (2004), Atkeson and Kehoe (2005), firm size is a key state variable for organizational restructuring or innovation. In these models, overall size is a proxy for firm productivity for accumulated ideas as in Klette and Kortum (2004) and will therefore influence innovation or optimal organization. In this class of models, two firms of the same size should optimally choose the same organization or make the same innovation decisions, even if they have different ages. In contrast, in models of "passive learning" such as Jovanovic (1982) or Acemoglu et al. (2007), firm age plays a more direct role beyond its influence on size. In these models, older firms optimally learn about parameters of their own production functions so that two firms of equal current size but different age might choose different organizational forms or make different innovation choices.

This distinction of size and age effects shapes our theory discussion below and guides our empirical strategy. To fix ideas we develop two very simple baseline models that are related to the two broad groups of mechanisms we are interested in: organizational restructuring and innovation. In particular, we outline size and age effects for a model of organizational capital, thereby relating organizational design choices over the life-cycle to static production functions. Furthermore, we build intuition for age and size effects in an endogenous innovation model, thereby relating the innovation life-cycle to dynamic or innovation production functions.

2.1 Firm Organization and Management Practices

Consider the following production function for firm i that is using labor L_{it} and organizational capital M_{it} to produce real output Y_{it},

$$Y_{it} = A_i L_{it}^{1-\alpha} M_{it}^{\alpha},$$

where A_i is initial or permanent productivity. Aggregate production is assumed to be given by a CES aggregator with constant elasticity of substitution η. The firm therefore faces the following maximization problem:

$$\begin{aligned} \max_{M_{it}, L_{it}} \; & P_{it} Y_{it} - \kappa_{it} M_{it} - w_t L_{it} \\ s.t. \quad & P_{it} = \left(\frac{Y_{it}}{Y_t} \right)^{-\frac{1}{\eta}} P_t, \end{aligned} \tag{1}$$

where the costs of organizational capital κ_{it} depend on firm-specific parameters. We use this cost specification to introduce passive learning for the costs of organizational capital in a similar fashion as in the Jovanovic model. There are underlying cost types

$$\bar{\varphi}_i \sim N(\mu_\varphi, \sigma_\varphi).$$

Each period, firms receive a noisy signal on these cost types, which are distorted by an iid noise term $\varepsilon_{it} \sim N(0, \sigma_\varepsilon)$,

$$\varphi_{it} = \bar{\varphi}_i + \varepsilon_{it}.$$

Priors on the underlying costs of organizational capital are given by

$$\varphi_{i0} = E_0[\varphi_i] \sim N(\mu_0, \sigma_0).$$

A firm receives a signal once a year during its lifetime so that its age also summarizes the number of signals a firm received. It can be shown that the posteriors of the optimal signal extraction problem evolve according to

$$\theta(a_{it}) = \left(\frac{\sigma_\varphi^2}{\frac{\sigma_\varepsilon^2}{a_{it}} + \sigma_\varphi^2}\right) \cdot \bar{\varphi}_i + \left(\frac{\sigma_\varepsilon^2}{\frac{\sigma_\varepsilon^2}{a_{it}} + \sigma_\varphi^2}\right) \cdot \frac{\varphi_{i0}}{a_{it}} + \left(\frac{\sigma_\varphi^2}{\frac{\sigma_\varepsilon^2}{a_{it}} + \sigma_\varphi^2}\right) \cdot \left(\frac{1}{a_{it}} \sum_{\tau=1}^{a_{it}} \varepsilon_\tau\right), \tag{2}$$

where a_{it} is the age of firm i at date t. For simplicity we assume that the costs of organizational capital are given by

$$\kappa_{it} = \exp\{-c \cdot \theta(a_{it})\} \tag{3}$$

Taking the first order conditions of (1) and plugging in for (3) then gives

$$\ln M_{it} \propto \ln L_{it} + c \cdot \theta(a_{it}). \tag{4}$$

Equation (4) provides a convenient summary of size and age effects in organizational restructuring over the life-cycle. The size effects simply reflect the fact that larger companies require more organizational capital to offset the diminishing returns of labor as a production factor. On the other hand, even conditional on these size effects, there will be age effects as firms learn about their idiosyncratic costs of organizational capital $\bar{\varphi}_i$ as their age a_{it} increases.

Note that although in this simple formulation of organizational capital, the size and age effects are log-linear, we will allow for a more general empirical specification that will be able to pick up non-monotonic or nonlinear size and age effects. We summarize the first main result based on our model of organizational capital as follows:

Result 1: Age and size effects of organizational capital

> Firms restructure their organization as they grow or as they age (i.e., adjust M_{it}). Hence, the model generates both conditional size and age effects: there are size effects conditional on firm age and age effects conditional on firm size.

Given this parsimonious baseline model of organizational capital we are now ready to discuss more specific mechanisms of organizational capital that are the focus of recent theoretical contributions. These more detailed mechanisms can be understood as providing micro-foundations of the accumulation of organizational capital in our baseline model. For example, a firm might add organizational layers to improve monitoring of employees as it grows larger thereby accumulating organizational capital and reducing the effect of diminishing returns to labor.

The following review of this recent literature is not intended to be comprehensive but will rather reflect the availability of direct measures of these theories in our firm-level survey data.

Organizational layers A natural source of diseconomies of scale is the business owner's time constraint as in Garicano (2000). In these models, a firm economizes on the business owner's time by taking advantage of the division of labor across organizational layers within the firm. Lower layers solve simple problems and report more complicated problems to higher layers, which in turn specialize in solving only complex problems. Such models of "knowledge hierarchies" or "management by exception" therefore offer a natural form of size effects whereby the number of organizational layers endogenously increases in overall firm size. Unconditional size effects have been documented in previous empirical work by Caliendo et al. (2015).[2] We extend their analysis along two dimensions: first, we estimate size

[2] Similar size effects can be generated in models in which managers are used to monitor employees to prevent them from shirking and in which additional managerial layers mitigate dis-economies of monitoring effort as managerial spans of control widen, see Calvo and Wellisz (1978).

effects conditional on firm age; second, we look at broader measures of organizational hierarchies.

Models of "management by exception" have also been used to generate age dynamics conditional on size. Garicano and Rossi-Hansberg (2012) describe a dynamic equilibrium model of exploration of new ideas and the exploitation of profits associated with existing ideas. When new technologies are first introduced or firms are founded, only the solution to the most common problems is known. Over time, as the organization faces more and more issues related to the commercialization of the new technology, managers learn more about complex problems and the firm adds more organizational layers. This addition of layers requires "time to build organizations," which is why the model in Garicano and Rossi-Hansberg (2012) introduces an age effect that is separate from the size effect. In a similar spirit, Rajan and Zingales (2001) present a model of optimal organizational layers of startups. Entrepreneurs hiring employees face a trade-off between expropriation risk of their ideas by employees and the need to induce employees to make firm-specific sunk investments to benefit from the division of labor. In their model, firms optimally start with few layers to minimize expropriation risk but add more layers over time as employees have made firm-specific investments that are sunk. In both Garicano and Rossi-Hansberg (2012) and Rajan and Zingales (2001) one should therefore expect to see more layers as firms age conditional on firm size.

Centralization A second form of organizational design choice that has received a lot of theoretical attention is decentralization. Conceptually, decentralization captures the intensive margin of organizational layers, since it is about the allocation of decision tasks across a given set of layers. In this context, as a firm grows larger, employees hired by the business owner are likely to acquire knowledge during the conduct of day-to-day operations delegated to them. Since the business owner does not naturally have access to the same knowledge, decisions in larger firms can be based on worse information. To exploit this information, firms can decentralize decision making as emphasized by Aghion and Tirole (1997), thereby trading off the loss of control by the business owner against the utilization of information acquired by their employees. Because this type of information might become more important as firms get larger, the prediction would be that firms tend to decentralize as they grow in size. In contrast, recent work by Akcigit et al. (2014) focuses on the opposite side of this trade-off. As firms get larger, employees need to be monitored since they otherwise abuse their decision authority for personal gain. Akcigit et al. (2014) show that this leads to the prediction that firms should optimally centralize decision making as they grow. Both Aghion and Tirole (1997) and Akcigit et al. (2014) therefore offer examples of firm size effects of decentralization, albeit with different predictions.

Acemoglu et al. (2007) formalize a variation of Aghion and Tirole (1997) that generates firm age effects. In their model, decisions can be decentralized to employees who make self-serving decisions with some probability that are unprofitable for the firm. However, the firm can learn over time what the optimal choice is, either from its own history or from competitors' choices. As a consequence, older firms are more likely to centralize decision making to exploit this learning effect while younger firms typically decentralize decisions.

Span of control A third organizational dimension that is closely related to layers and decentralization is the span of control, defined as number of employees reporting to a given manager. An example of a model along those lines that generates firm size effects is Calvo and Wellisz (1978), which was previously discussed in the context of models generating a size effect in the overall number of organizational layers. In the Calvo-Wellisz model, firms expand first by adding employees and monitoring them more intensively with a given number of managers. In other words, managerial span of control increases as a firm expands. Eventually, the addition of new layers limits the increase in the span of control. As a result, the Calvo-Wellisz model predicts a slowdown in the increase of the span of control as new layers are added.

In contrast, the model of Rajan and Zingales (2001) generates firm age effects in the span of control. Here, young firms that strive to minimize expropriation of ideas start out with few managers per employee and therefore a high span of control. As the firm ages and employees make firm-specific sunk investments that reduce expropriation risk, more employees per manager are higher to exploit the division of labor. As a result, the span of control decreases as the firm ages. This is again closely connected to the overall number of layers, since the span of control falls at the same time as the overall number of layers increases.

Performance pay As previously discussed, many models of firm organization formalize the intuition that employees cannot be perfectly monitored all the time, thereby creating moral hazard problems. To counter these moral hazard problems, firms can adopt performance pay to counter shirking or to reduce selfish decision making. From this perspective, larger firms might naturally be subject to more moral hazard problems, since it is harder for business owners to monitor a large number of employees. A first prediction might therefore be that larger firms are more likely to adopt performance pay, which is thus a size effect, even conditional on firm age.

Manso (2011) on the other hand provides a rationale for firm age effects in this context. In his model, firms pursuing exploration of new ideas instead of exploitation of existing opportunities optimally do not adopt performance pay based on current profits but instead use an incentive system that allows for "tolerance for failure." As a result, one would expect that as firms age and shift from exploration of new ideas towards exploitation of existing profit opportunities, they are more likely to adopt performance pay, an example of a firm age effect.

Alternatively, the composition of tasks conducted by employees might change over time. Harder-to-monitor tasks such as innovation and creative activities as in Holmstrom (1989) might be more important for younger firms, while older firms might focus more on relatively standardized tasks associated with the production and distribution of their established products. In this alternative view, performance pay adoption is more likely for younger firms than for older firms. Since the task composition in this theory is a function of firm age and not primarily of firm size, the prediction that younger firms might be more likely to adopt performance pay is a pure age effect.

2.2 Innovation

In this section we build a baseline model that captures the difference between firm age and firm size effects on innovation, similar to the approach in the previous section. The starting point is the endogenous innovation model by Klette and Kortum (2004) with the following innovation production function:

$$Z_{it} = \left(R_{it} \cdot \exp\{\theta(a_{it})\}\right)^{\frac{1}{\gamma}} N_{it}^{1-\frac{1}{\gamma}}, \tag{5}$$

where Z_{it} is the flow rate of innovation and R_{it} are innovation inputs such as R&D and the number of scientists. In this model, firm i is interpreted to be a collection of products j for which the firm has acquired the frontier technology through innovation. Therefore, N_{it} is the number of products for which the firm currently has the frontier technology. This is proportional to firm size.[3] We note that in this

[3]We follow Aghion, Akcigit and Howitt (2013) in specifying the underlying economic environment by assuming an aggregate production function that combines products j into a composite good Y,

$$\ln Y_t = \int_0^1 \ln y_{j,t} dj$$

with product-specific static production functions given by $y_{jt} = A_{jt} l_{jt}$. The economy is assumed to be populated by a measure of workers L that can either work in production or in R&D in incumbent firms or entrants.

model, static profits per product j are proportional to overall market size Y_t, i.e., $\pi_{jt} = \pi \cdot Y_t$, where π is a constant. Since all products enter in a symmetric way, we can write the overall flow profit of firm i as $\pi_{it} = \pi_t \cdot N_{it}$.

As before $\theta(a_{it})$ can be understood as capturing the posterior of an uncertain cost parameter as given by (2). The main difference is that here the uncertainty is about a parameter of the innovation cost function, so that firms passively learn about their comparative advantage in generating innovations. Defining $z_{it} = \frac{Z_{it}}{N_{it}}$, the value function can be written as

$$\begin{aligned} rV_{it}(N_{it}) - \dot{V}_{it}(N_{it}) = \max_{z_{it}} \quad & N_{it}\pi_t - w_t \exp\{-\theta(a_{it})\} N_{it} z_{it}^{\gamma} \\ & + N_{it} z_{it} \left[V_{it}(N_{it}+1) - V_{it}(N_{it}) \right] \\ & + N_{it} \chi_{it} \left[V_{it}(N_{it}-1) - V_{it}(N_{it}) \right] \end{aligned}$$

where χ_{it} is the probability that a competing firm innovates and steals a specific product from firm i. This dynamic value function adds to the static profits the costs of innovation inputs as well as potential capital gains and losses from either acquiring new products or losing products to competitors.

Solving for the optimal value function and then substituting out R&D inputs R_{it} and using the fact that $\ln L_{it} \propto \ln N_{it} + \ln \frac{Y_t}{w_t}$, it can be shown that the innovation production function in equilibrium can be rewritten as

$$\ln Z_{it} \propto \frac{1}{\gamma - 1} \theta(a_{it}) + \ln L_{it}. \tag{6}$$

Equation (6) clarifies the sources of firm age and firm size effects in innovation. As in Klette and Kortum (2004), larger firms can build on more ideas and are therefore more likely to innovate. In contrast, our passive learning extension of the Klette-Kortum model generates an additional age effect. As firms age, they learn about their underlying comparative advantage in innovation so that innovation might become more likely for older firms, conditional on size. We summarize the second main result based on our model of endogenous innovation as follows:

Result 2: Age and size effects in innovation

> Firm innovation flow rate Z_{it} is a function of firm size and firm age in our extended version of the Klette-Kortum model. The model generates both conditional size and age effects: there are size effects conditional on firm age and age effects conditional on firm size.

Starting with work by Klette and Kortum (2004), endogenous innovation models generate realistic firm dynamics by combining innovation by entrants with innovation by incumbents. In the model of Klette and Kortum (2004) without our extension, if one would follow a cohort of firms as they age, older firms would be larger and more likely to innovate. However, our extension highlights that this is primarily a size effect, so that conditional on size there should not be any age effect, see the discussion in Klette and Kortum (2004). Another important observation is that the size effects in Klette and Kortum (2004) is monotonic in size, so that larger firms are always more likely to innovate than smaller firms. Our empirical findings will highlight this monotonicity prediction and contrast different results for manufacturing vs. services firms.

Incremental innovations Akcigit and Kerr (2010) offer an extension of the Klette and Kortum model in which young and small firms explore radical innovations while older and larger firms concentrate on non-radical innovations that focus on exploiting existing profit opportunities. In contrast to Klette and Kortum (2004), the model by Akcigit and Kerr makes predictions about the composition of the types of innovations generated by firms. As a result, the likelihood to generate radical innovations should decline with firm age and firm size, while the likelihood to generate non-radical, incremental innovations should increase. It should be noted that these predictions are primarily driven by size effects, similar to the effects in Klette and Kortum (2004). That is, the reason younger firms tend to be more likely to pursue radical innovations is related to their size rather than the fact that they are not very old.

Product or process innovations Innovations in Klette and Kortum (2004) and related models can be understood either as process innovations that imply a reduction in marginal production costs or as quality improvements of existing products. While these models are typically silent about the difference between product and process innovations, Klepper (1996) provides a theory of firm age dynamics that explicitly models this dimension. Younger and smaller firms are more likely to generate product innovations, while older and larger firms focus on process innovations. Klepper (1996) therefore provides additional predictions about the type of innovations likely to result as firms age. As in the case of endogenous innovation models discussed before, the theory of Klepper (1996) works primarily through firm size instead of firm age effects. The reason is that profits from process innovations scale with firm size, while profits from product innovations do not. This is the main reason older and larger firms specialize in process innovations, while smaller and younger firms specialize in product innovations.

3 Data and Methodology

In this section we describe our dataset and the methodology we use to estimate life-cycle dynamics of establishments and firms.

3.1 Data

To study the dynamics of firm organization and innovation activities, we use the Workplace and Employee Survey (WES), a random stratified establishment-level panel conducted by Statistics Canada and drawn from the universe of all Canadian firms, including non-manufacturing industries. The survey has a cross-sectional dimension of approximately 6,500 establishments per year over the time period from 1999 to 2006. The data is drawn from a stratified sampling frame, which allows the construction of sampling weights, making the data representative of around 1 million private employer establishments in the Canadian economy. We focus on the sub-sample of around 5500 for-profit business firms. As in other government-sponsored surveys, response to the WES was mandatory so that the overall response rate was typically close to 90% in all survey years.

A particularly attractive feature of the implementation of the WES is that Statistics Canada invested considerable effort in ensuring the precise measurement of firm exit decisions. In other establishment-level surveys—especially in the developing world—it is often unclear whether sample attrition is driven by firm exit or by non-response. For the WES, Statistics Canada instead followed up on non-respondents even up to a year later to find out whether the firm had indeed gone out of business or failed to respond. If the firm had not gone out of business, every effort was made to convert non-respondents to respondents, which explains the very low attrition rate in the survey.

Importantly, the WES provides various measures of firm organization and management practices as well as detailed measures of organizational change and innovation activity. The WES has at least three distinct advantages compared to currently available empirical work on firm age dynamics, such as Foster, Haltiwanger and Syverson (2013) and Hsieh and Klenow (2014). First, our combined dataset is representative of the entire Canadian economy, including the large service industry, rather than being limited to a few sectors such as manufacturing. Since the major part of economic activity in advanced economies such as Canada or the US takes place outside of manufacturing, we consider this a major advantage. Second, we can rely on birth year information that was directly asked in the survey and that we checked for consistency and validated with similar age information from the Canadian Annual Survey of Manufacturing. Hence, despite the fact that the WES was conducted for only eight consecutive years, it provides reliable firm age similar to the measures available in Census-based studies that draw on much longer panels, such as Foster et al. (2013) and Hsieh and Klenow (2014). Third, the WES includes information on whether establishments surveyed are single-unit firms or are part of existing multi-unit firms. Since the survey was conduced at the establishment level, we will distinguish between firms and establishments where necessary but will use the terms firms and establishments interchangeably in the context of single-unit firms. The difference between single-unit firms and establishments of existing multi-unit firms turns out to be crucial for the empirical analysis, as it allows us to separately track life-cycle dynamics of startups and of new establishments of existing firms. Although most theories of firm dynamics we discussed do not make a clear distinction between establishment and firm dynamics, we will see that empirically there are large differences.

3.2 Methodology

For the analysis of life-cycle patterns it is important to separate age, cohort and time effects due to the fact that age can be recovered from a combination of cohort (or birth year) and time. We would like to stress from the outset the well-known issue that for our object of interest—namely age effects—panel data alone does not necessarily address key identification issues such as the separation of age, cohort and time effects; see e.g., Hall, Mairesse and Turner (2007) and Schulhofer-Wohl (2013).

In earlier versions of this paper we experimented with different ways to separate age from cohort and time effects, such as the methodology developed by Deaton (1997). A more conservative and in our opinion more transparent specification however only uses within-establishment variation to calculate age effects, which is why this is the specification we chose. All our results are robust to using a more sophisticated approach such as the one developed by Deaton (1997). Note in particular that establishment cohort is a time-invariant feature of each establishment, so that it will be absorbed in the establishment fixed effect. The use of establishment fixed effects will go beyond controlling for cohort effects by absorbing all time-invariant cross-sectional variation across establishments. This specification has therefore the added advantage that it facilitates the interpretation of our results as differences in organization and innovation that are correlated with within-firm growth as opposed to cross-sectional variation in firm size.

Another feature of our methodology is that we non-parametrically estimate outcomes as a function of firm age and firm size while controlling for a full set of establishment and time fixed effects. This allows us to identify possible non-monotonic size or age effects, which turns out to be empirically relevant.

Firm and establishment size dynamics To characterize establishment growth over the life-cycle, we decompose log-size $\ln(s_{e,\{a,c,t\}})$ of establishment e that is a years old at time t and is therefore part of cohort c as

$$\ln\left(s_{e,\{a,c,t\}}\right) = \lambda_a + \tau_t + \xi_c + \xi_e + \texttt{error}. \tag{7}$$

In words, we decompose establishment size into:

- conditional age effects λ_a that depend on the age of the establishment. We non-parametrically estimate the shape of the age profile by creating 20 age dummies for each age from 1 to 19 as well as "20+". We chose to pool of age effects for establishments 20 years and older because we previously estimated a full age profile from ages 1 to 50 and found that age dynamics slow down after the age of 20. The dummy for age 0 is omitted so that all estimates should be interpreted to display age effects relative to age 0 establishments, some of which are startups others are new establishments of already existing parent firms.

- time effects τ_t that affect all establishments the same way.
- cohort effects ξ_c that effect all establishments of the same cohort the same way.
- establishment fixed effects ξ_e that captures time-invariant features of a specific establishment.

Our preferred measure of size is the number of employees since this measure allows us to easily compare our results to patterns documented by previous empirical studies. Based on the separability of firm level heterogeneity assumed in (7), we estimate the following specification:

$$\ln(s_{e,\{a,c,t\}}) = \nu + \sum_{a=1}^{A} \lambda_a \cdot D_a + \sum_{t=1}^{T} \tau_t \cdot D_t + \sum_{e=1}^{E_t} \xi_e \cdot D_e + \texttt{error}, \tag{8}$$

where D_a, D_t are age and time dummies and D_e represent a full set of establishment fixed effects with E_t as number of establishments.

Organizational restructuring and innovation For the estimation of the life-cycle dynamics of management practices and innovation, we keep the basic strategy outlined before, but add size effects. We start with a similar micro-level decomposition:

$$\ln\left(x_{e,\{a,c,t,s\}}\right) = \kappa_s + \lambda_a + \tau_t + \xi_c + \xi_e + \texttt{error}, \tag{9}$$

where the outcome $x_{e,\{a,c,t,s\}}$—which is now an organizational practice or innovation type—is a function of age effects λ_a, time effects τ_t, cohort effects ξ_c and establishment fixed effects ξ_e. Hence, in order to separate age effects from size effects, we add

- conditional size effects κ_s, which we capture non-parametrically by creating 15 different firm size bins ranging from "5-9 employees" to "1000+ employees". The smallest size category "1-4 employees" is omitted so that all size effects should be interpreted as being relative to establishments with 1 to 4 employees.

We therefore estimate the following specification for life-cycle dynamics of organizational restructuring and innovation:

$$\ln(x_{e,\{a,c,t,s\}}) = \nu + \sum_{s=1}^{S} \kappa_s \cdot D_s + \sum_{a=1}^{A} \lambda_a \cdot D_a + \sum_{t=1}^{T} \tau_t \cdot D_t + \sum_{e=1}^{E_t} \xi_e \cdot D_e + \texttt{error}, \quad (10)$$

where we added a set of size bin dummies D_s. In other words, the age and size effects of organizational restructuring and innovation that we will be reporting are conditional effects: age effects controlling for establishment size, and establishment size effects controlling for age. We focus on the conditional results for the sake of brevity. In unreported results for this paper we also estimated specifications with either only size or only age effects without any major differences in results.

Finally, due to confidentiality issues, some estimates of size effects have not been cleared by Statistics Canada. We mark these estimates in the figures by reporting a value of zero and collapsing the standard errors to zero as well. However, for these suppressed estimates, Statistics Canada did provide the sign as well as information on whether these estimates were statistically significant. Whenever needed we will report these qualitative results, too.

4 Results

In this section we describe our main results of the estimated establishment-level life-cycle dynamics of firm growth, organizational (re)structuring, and innovation.

4.1 Firm Growth over the Life-Cycle

General life-cycle patterns We start by reporting life-cycle dynamics estimated using (8) summarized in the left panels of figure 1. The top panel shows age effects for the entire (aggregate) economy, except retail and wholesale. We exclude these sectors because firm growth in these sectors is mainly driven by expanding the number of establishment instead of growth within an establishment. We contrast these economy-wide life-cycle dynamics with the age effects in two sectors that will be the main focus of our analysis: manufacturing and services. For the Canadian economy, this means that the main sector that drives the differences between economy-wide average outcomes and manufacturing or service specific patterns is mostly mining and oil or gas extraction sectors.

A number of notable features emerge by comparing these different sectors. First, both manufacturing and services exhibit steeper age effects than the economy as a whole, reflecting the fact that age dynamics in the mining and oil or gas extraction sector are more muted. Second, the age profile in manufacturing is steeper than in services in the first 10 years but at the same time flattens out earlier, too. Specifically, 10 year old manufacturing establishments are 50% bigger than at entry, while 10 year old service establishments are only about 23% bigger than at entry. However, 20 year old service establishments are more than 50% larger than at entry, while manufacturing plants at age 20 are barely larger than at age 10.

We can put these results in an international perspective for the manufacturing sectors, for which Hsieh and Klenow (2014) provide results for the US, Mexico and India. However, it is worth noting that Hsieh and Klenow (2014) do not exclusively use within-establishment variation. Keeping this caveat in mind, we start with the US comparison. The age advantage of 10-to-14-year-old manufacturing plants in the US is 2.4, which is substantially larger than the Canadian manufacturing plants over the same age interval. In contrast, Indian manufacturing plants barely grow, and even if they survive until the age of 30-34, are only around 40% bigger than the typical entrant. Mexican plants, on the other hand, behave similarly in their age dynamics to Canadian manufacturing plants. By age 10-14, the representative surviving Mexican plant is about 40% bigger than an entrant, compared to 50% in Canada. Hsieh and Klenow (2014) also report age effects for an even broader set of countries, showing that most countries are somewhere between the extreme cases of India and the US in terms of their age dynamics. From this perspective, an analysis of the determinants of age dynamics in the Canadian economy could be considered quite representative of the mechanisms of such age dynamics in countries, other than the US and India.

The role of selection How does selection affect life-cycle dynamics? We are particularly interested in this question as there are at least two ways in which one can think of the influence of selection on life-cycle dynamics. First, strongly growing establishments might be more likely to survive so that much of the life-cycle dynamics might be driven by the dynamics of these surviving establishments. Second, life-cycle patterns might be driven more by the fact that shrinking establishments exit. Fortunately, with our data it is possible to gauge the effect of selection by comparing the life-cycle dynamics of all establishment with the life-cycle dynamics of continuing establishments.

The right panels of figure 1 present the age dynamics of continuing establishments only. Comparing these with the left panels it becomes clear that although the removal of exiting establishments does influence the estimated life-cycle patterns slightly, most of the age dynamics are indeed driven by the growth of continuing establishments. These patterns are another manifestation of the "up or out" dynamics documented for US firms by Haltiwanger et al. (2013).

Comparing startups with new establishments of existing parent firms An important question that will help to delimit the domain of our empirical analysis is whether the establishment life-cycle patterns we documented are driven by the birth of new firms (i.e., by startups) or by newly created establishments of already existing multi-unit firms. This question is of particular interest as most theories we outlined in section 2 do not explicitly distinguish between firms and establishments. Stated differently, although it is plausible that life-cycle dynamics are driven by a new firm learning about how to organize a new business, it is similarly plausible that there are establishment-level learning effects, such as an existing business learning to adapt to a new geography or to optimally adapt its organization to a new set of employees.

To answer this question, in figure 2 we separately estimate (8) for startups (left panel) and new establishments of existing multi-unit firms (right panel). Figure 2 shows that while the life-cycle patterns of of new firms (i.e., following startups) are comparable to the overall life-cycle patterns of continuing establishments, the life-cycle patterns of continuing establishments that are part of multi-unit firms are very different. The life-cycle pattern of new establishments that are part of an already existing multi-unit firm are mostly flat for establishments in service industries and if anything are even decreasing for manufacturing establishments.

These results show that most of the life-cycle dynamics are driven by startups or new firms instead of new establishments that are part of pre-existing multi-unit firms. This finding is reminiscent of recent evidence in the literature on the age dynamics of multinational subsidiaries, such as Garetto, Oldenski and Ramondo (2016). This literature shows that subsidiaries of multinational companies have flat life-cycle patterns, once one controls for growth of the home country (parent) firm.

Based on these results, in what follows we will focus on the organizational restructuring and innovation activities of startups.

4.2 Age vs. Size Effects

In this section we contrast firm age effects of organizational restructuring and innovation to firm size effects, for the sample of startup firms (hence firm and establishment effects are identical). We start each subsection with a brief discussion of the measures used in the empirical analysis that are informed by our theory discussion in section 2 and then move on to discuss the results. All reported results are estimated using specification (10).

To conserve space, we focus on the results in manufacturing in this section and will then use section 4.3 to point out the major differences between service and manufacturing sectors.

4.2.1 Organizational Restructuring

Organizational layers We use two alternative measures to proxy for organizational layers. First, we follow Caliendo et al. (2015) in using the number of different occupations present in a firm to proxy for the number of layers. The WES provides detailed information on whether an establishment's employees fall into the the following categories: managers, professionals (employees with a university degree), technical/trades, marketing/sales, clerical/administrative, production workers without trade/certification, and "other occupations." Using occupations to determine the number of layers has the advantage that it enables us to compare our empirical results with related previous studies in this literature. We will refer to these measures as "occupational layers."

Second, the WES includes detailed information regarding real decision authority on tasks across layers in the organizational hierarchy, which is the basis for an alternative measures of "decision layers." In contrast to formal reporting patterns summarized in occupational titles or organizational charts, real decision authority avoids measurement issues that emerge if managers with formal authority merely "rubberstamp" decisions of subordinates, a point forcefully made by Aghion and Tirole (1997). The survey questions are similar to measures of worker autonomy used in Bresnahan, Brynjolfsson and Hitt (2002) and Bloom, Garicano, Sadun and Van Reenen (2013) in that they allow us to measure to what degree decisions across 12 potential tasks are made by principals or agents. Specifically, the survey question is, "Who normally makes decisions with respect to the following activities?" The respondent is then given a choice of 12 possible activities, from "Daily planning of individual work" to "Quality control" to "Product and service development." There are six possible responses to the question of who makes decisions: non-managerial employees, work group, work supervisor, senior manager, individual/group outside the workplace—typically headquarters for multi-establishment firms—and business owners.

Our measure of decision layers counts the number of layers that are involved in any decision task among the 12 possible tasks in the survey. If a decision layer is not involved in any decision among those tasks, we infer that it is effectively not present. This strategy for measuring organizational layers has both advantages and disadvantages. One possible disadvantage is that our measure provides only a lower bound on the number of decision layers in the organizational hierarchy. Additional layers, which effectively do not play any role in decision-making even for very routine tasks such as "daily planning of work" could in principle exist in the organization and would be observed in say an organizational chart. At the same time, we also view this potential disadvantage as a unique advantage of our measure of layers. Specifically, additional layers with no decision authority among any relevant tasks are likely just a result of office politics and reflect inefficient growth of bureaucracy and hence are not an indicator of the division of labor within the organization. While the growth of inefficient bureaucracy is an interesting additional dimension to consider in future work, in this paper we focus on recent theories of knowledge hierarchies and the division of labor within organizations.

The first four graphs of figure 3 display the results of estimating (conditional) firm age and firm size effects using equation (10). Interestingly, there are no clear age effects irrespective of the measure for organizational layers we use. This stands in sharp contrast to the clear size effects shown in the right-side panels of figure 3. Our results on occupational layers are consistent with similar estimates by Caliendo et al. (2015), who show for French firms that an increase in value added is correlated with an increase in the number of occupational layers.

Regarding firm size effects, a number of novel patterns emerge. First, occupation patterns roughly follow an inverted U-shape pattern, where the number of occupational layers first increases and subsequently falls as manufacturing firms grow past 200 employees. This non-monotonic size effect is neither predicted by models of knowledge hierarchies such as Garicano (2000) nor by monitoring hierarchy models in the spirit of Calvo and Wellisz (1978). Second, we do see differences in size effects of occupational and decision layers. While the number of occupational layers grows over several size classes, most of the changes in the number of decision layers occur while firms are still relatively small. These differences seem to suggest that decision layers are quite a different organizational phenomenon compared to occupational layers. Specifically, restructuring of decision layers systematically happens while firms are still quite small, but is less important when firms reach a certain size. In contrast, even relatively large firms past 100 employees are still involved in restructuring the overall number of occupational layers.

Centralization Our measure of centralization is based on the same data described above for the construction of decision layers. However, here we exploit more of the variation in the number of tasks decided by different organizational layers. Specifically, we separate decision layers into "managers" on the one hand, which include business owners, senior management and work supervisors, and "non-managerial employees" on the other hand. We focus on the difference between "managers" and "non-managers" for several reasons. First, data that allows us to differentiate between senior management and business owners is only available for around half of the sample. Second, the model of Akcigit et al. (2014) seems to apply equally for the difference between managers and non-managerial employees, although it is stated in terms of business owners vs. managers.

Given this difference between managers and non-managerial employees, our measure of centralization counts the number of tasks that are exclusively completed by managers without any decision-making by non-managers. While our data in principle allows us to measure joint decision-making by managers and non-managerial employees, we instead focus on exclusive decision-making in order to make the results easier to interpret.

As the last two graphs of figure 3 show, we are not able to document any clear systematic pattern of the centralization of tasks, neither for age nor for size effects. Note, that one caveat to this conclusion is that estimates for size effects for the size classes "5-9", "15-19" and "750-1000" have been suppressed by Statistics Canada for confidentiality reasons. However, Statistics Canada did allow us to report that these size effects estimates in all of these cases were insignificant. To summarize, we fail to find evidence for systematic changes in centralization with age and size for manufacturing firms.

Span of control As described before, the WES provides detailed data on occupations of employees. We use this data to calculate the inverse of span of control, which is the percentage of employees that are managers.

The top panel of figure 4 shows the profile of firm size and firm age effects for manager per employee based on (10). As before, we fail to establish evidence of systematic age effects, once firm size is controlled for. On the other hand, firm growth seems to correlated with changes in the span of control. Specifically, as firms grow, the number of managers per employee falls, or conversely, the span of control for each manager systematically increases. These firm size effects of managerial span of control warrant several comments. First, the overall patterns by themselves seem consistent with models of monitoring hierarchies, such as Calvo and Wellisz (1978). Second, the span of control increases until manufacturing firms reach a size of about 200 employees, but stays flat afterwards. This is notable, as monitoring hierarchy model such as Calvo and Wellisz (1978) connect changes in the span of control with the overall number of layers. Specifically, an increase in the number of layers should be correlated with a slow down in the increase of span of control. The logic is that adding more layers limits the number of subordinates each manager has to monitor. Instead we observe that growing firms exhibit both a strong increase in the number of occupational layers as well as an increase in the span of control. This relation is even less clear if we use the number of decision layers as a measure of the span of control, as this measure of layers does not exhibit size effects beyond firms with less than 20 employees. Third, current models of monitoring hierarchies are unable to explain why size effects of span of control are basically flat for very large firms as shown in figure 4.

Performance pay The WES survey data offers a variety of information on performance-based compensation used by firms. Specifically, it allows us to measure four different types of performance pay: individual incentive pay such as bonuses, commissions, piece-rates etc.; group or team incentives; profit sharing agreements; and stock-based compensation. Standard principal-agent theory typically characterizes very general forms of state-contingent compensation contracts to solve the moral hazard problem. Consequently, we measure the presence of performance pay with an indicator that is one if any form of performance pay is present. We exclude stock compensation from this measure, since information on stock compensation is completely missing for one year and only a very small fraction of firms offer stock compensation to their employees. For more details on the performance pay data, see our complementary analysis in Hong, Kueng and Yang (2016).

We display the estimated age and size effects for performance pay adoption in the bottom panel of figure 4. As with other measures of organizational practices, there are no discernible age effects in the adoption of performance pay. Perhaps surprisingly, manufacturing firms do not systematically adopt performance pay measures as they grow. However, this result only applies to manufacturing firm, while firms in service industries to increasingly adopt performance pay measures as they grow, a point we will return to in section 4.3 where we contrast the size effects of manufacturing and service firms.

4.2.2 Innovation

We use survey measures of innovative activities included in the WES to inform recent theories of endogenous innovation. These measures are constructed following the Oslo-manual methodology outlined in OECD, Eurostat and EC (1992). Specifically, the WES asks respondents to report four possible types of innovations: new products or services, new processes, improved products or services, and improved processes. The survey form also asks establishments to categorize innovations that "differ significantly in character" from products or processes used before as "new." In this sense, the difference between "new" and "improved" could be interpreted as capturing the difference between drastic and incremental innovations.

An important question is whether these subjectively reported innovation measures are reliably related to real firm performance or whether the measured variation is potentially spurious. To address this question, it is worthwhile to point out that survey data on innovation following the Oslo-manual has been used in a variety of econometric studies, see Mairesse and Mohnen (2010). In fact, there exists a long literature on the performance effects of these innovation measures, such as Van Reenen (1997), Garcia, Jaumandreu and Rodriguez (2005), Peters (2008). These studies combine such subjective innovation data with different econometric techniques such as instrumental variables or structural estimation and analyze data from a variety of different countries such as the UK, Germany, Italy and various other European countries. A common finding across studies is strong positive employment growth effects of product innovations at the firm level, while process innovations mostly result in weakly negative employment growth. We additionally validate these innovation measures, by showing that innovation by other firms in the industry significantly depresses growth at non-innovating firms.[4]

Product and process innovation To measure product and process innovation activity we count all types of reported innovations, whether they are "new" or "improved" and categorize them into product and process innovations.

Figure 5 shows conditional age and size effects for innovation activities based on (10), where the left panels show conditional firm age effects and the right panels present conditional firm size effects. As was the case for all previous measures of organizational restructuring we analyzed in section 4.2.1, we fail to find systematic evidence for age effects conditional on firm size.[5]

In contrast, there are clear size effects for all types of innovation. These results are notable from the vantage point of several theories of product and process innovation. First, as firms grow larger, they are more likely to innovate. These conditional firm size effects are exactly what endogenous innovation models such as Klette and Kortum (2004) would predict. Second, although the result that the probability of

[4]These additional results are available upon request.

[5]Results on unconditional age effects are almost identical and available upon request.

process innovation increases with firm size is consistent with the model of Klepper (1996), the increase in the probability of product innovation is not. Third, although conditional firm size profile of innovation probabilities has a positive slope, there is also strong evidence of curvature as (very) large firms run into diminishing returns to innovation.

Incremental innovations To measure incremental innovations we count innovations whenever they are classified as "improvements," irrespective of whether they are product or process innovations. The bottom panel of figure 5 shows the conditional age and firm size profiles for incremental innovations, which are similar to the patterns for product and process innovation. In particular, recall from our theory discussion that models such as Akcigit and Kerr (2010) predict that the probability of incremental innovation rises with firm size, a feature we find confirmed in the conditional size effects in the bottom right graph of figure 5. In contrast, there is again no evidence of systematic age effects once we control for firm size.

One slight difference between the conditional firm size profile of incremental innovations and the profiles for product and process innovations is that incremental innovations do not seem to run into similar diminishing returns as the latter two types of innovations. This might be due to the fact that larger companies that run into diminishing returns when working on new innovations (i.e. all the low-hanging fruit has been picked) shift their focus to improve their existing product lines and services, which shows up in increase incremental innovations. This is consistent with the mechanism described in Garicano and Rossi-Hansberg (2012), which in this context shows up in the evolution of different types of innovations as a function of firm size instead of organizational layers.

4.2.3 Summary: Conditional Size and Age Effects

Building on our theory discussion of conditional age and size effects in section 2, we asked whether organizational restructuring and innovation are mainly driven by firm age or by firm size effects. We especially focused on the firm dynamics of startups (i.e., new single-unit firms) as we showed in section 3 that it is mostly their life-cycle dynamics that is key to understanding overall life-cycle dynamics.

Using a variety of measures of management practices, firm organization, and innovation activities we found no evidence for age effects once we controlled for firm size. Although we mostly displayed results for manufacturing firms, this result extends to firms in service industries as well. In other words, for both manufacturing and services there is no evidence for passive learning models of organizational restructuring and innovation.

In contrast, our results are broadly consistent with the existence of substantial conditional firm size effects. In other words, startups systematically restructure some of their organizational practices as they grow, and larger firms become more likely to innovate. In the next section we will therefore focus on these conditional firm size effects and contrast our results on size effects that so far were based on the sample of manufacturing firms with the size effects obtained for the sample of firms in service industries.

4.3 Comparing Size Effects in Manufacturing and Services Startups

In this section we compare size effects for manufacturing startups to size effects for service startups. There are several reasons we are interested in this particular sectoral comparison. First, the service sector dominates economic activity in many advanced economies, hence it is interesting in its own right, especially in light of the fact that many empirical studies of firm- or establishment-level data are restricted to the manufacturing sector. Second, service sectors are often perceived as being less capital intensive and more labor intensive. As a result, organizational restructuring intended to offset diminishing returns from labor inputs might be more important for service firms than for manufacturing firms. Third, since manufacturing centers around physical products, constant product or process improvements might be more important than in service industries. Manufacturing firms might therefore exhibit different size effects than service firms.

4.3.1 Organizational Restructuring

Organizational layers The top and middle panel of figure 6 compare the size profiles of the number of layers for manufacturing firms (left) and firms in service industries (right). Overall, the conditional size effects display an interesting degree of similarity with some important differences. Service firms display a similar inverted U-shaped patterns in occupational layers as manufacturing. However, the reduction of occupational layers at (very) large firms is much more pronounced for service firms. While our size effect estimates suggest that the number of occupational layers of a firm with more than 1000 employees is still higher than for firms with 1-4 employees for manufacturing firms, the same is not true for service firms. In fact, the number of occupational layers for service firms with more than 1000 employees is not statistically distinguishable from firms of size 1-4. We are not aware of any theory that is able to explain the difference in size effects of occupational layers of service firms compared to manufacturing firms.

The inverted U-shape pattern of the number of layers in size is also more pronounced for service firms in our measure of decision layers. While for manufacturing firms most of the change in decision layers is concentrated at the smallest firms, our estimates suggest that changes in decision layers for medium-sized firms are common in service industries.

Centralization The bottom panel of figure 6 summarizes our comparison of centralization size effects in manufacturing vs. service firms. Due to confidentiality issues, the firm size effects on centralization for the largest service firms have been suppressed by Statistics Canada. However, Statistics Canada allows us to report that the estimates are also positive and statistically significant at the 5% level. Hence, the profiles displayed are relatively smooth even in the right tail of the firm size distribution.

The bottom right figure shows that service firms display a tendency towards centralization of decision making as they become bigger. This pattern is consistent with the prediction of the incomplete contracting model of Akcigit et al. (2014), where firms optimally centralize as they grow in response to higher expropriation risk. The comparison between manufacturing and service firms also suggest that such a type of expropriation risk is more important for services than for manufacturing. This might reflect the fact that former employees setting up a competing firm face lower entry barriers in services where less capital is needed to compete with their former employers.

Performance pay Comparisons of performance pay adoption between firms in service industries and manufacturing firms are shown in the bottom panels of figure 7. As service firms become larger, they are more likely to adopt performance pay incentive schemes. This is especially true for very large service firms, which are significantly more likely to have some form of performance pay. This stands in stark contrast to the absence of size effects in performance pay adoption for firms in manufacturing industries.

The difference in adoption patterns of performance pay are potentially reflecting deeper differences between service provision and good production, as incentives pay for services such as consulting and other business services might be more important than for the typical manufacturing tasks. Another possible explanation for this difference is that service firms beyond a certain size might shift from exploration of new business ideas towards exploitation, therefore starting to reward based on current performance. This would be consistent with the model of Manso (2011). Manufacturing firms in contrast might keep innovating, even as they grow large and as a result might not be more likely to adopt pay based on current performance. We will therefore now turn to the analysis of size effects in innovation patterns.

4.3.2 Innovation

Figure 8 contrasts size effects in our three types of innovation for firms in service industries (right side) with the earlier results for manufacturing firms, repeated for convenience on the left side. We see that innovation patterns as a function of size are vastly different between manufacturing and service firms. While innovation probabilities in manufacturing monotonically increase in firm size, there is no such pattern for service firms. On the contrary, innovation exhibits very strong non-monotonic patterns. While firms with 15 to 50 employees become more innovative than the smallest firms (with less than 4 employees), innovation activities strongly decline as firms pass 150 employees. Seen through the lens of our baseline Klette-Kortum models, this suggests that innovation patterns for service firms display strong diseconomies in firm size. Note that in this context the similar and non-monotonic patterns apply not only to product/service innovations but also to process innovations and to incremental innovations. In other words, these results are not likely to be driven by the mere fact that service firms set up their business model once and then keep improving their processes. Instead, our results suggest that improvement in processes slows down for large firms as well.

These stylized facts do not only have potentially consequences for modeling innovation dynamics in service industries, but are also suggestive of important aggregate consequences. For example, sectoral shifts from manufacturing towards services might contribute to a slowdown in aggregate innovative activity and therefore productivity growth. In this sense, our innovation results are complementary to recent work by Bloom, Jones, Van Reenen and Webb (2016) who compile evidence across various industries that the number of researchers to generate a given amount of innovation as disproportionately increased over time.

5 Conclusion

Based on recent theories of management practices and organizational capital as well as endogenous innovation, we empirically analyze the mechanisms behind firm life-cycle dynamics. We establish several key novel stylized facts. First, we show that firm life-cycle dynamics are driven by startup dynamics, which is the reason we focus on life-cycle growth patterns of startups.

Second, for these startup life-cycle dynamics, we document a number of systematic patterns that shed light on the mechanisms driving startup growth, differentiating size and age effects. Broadly, we find evidence for systematic restructuring of management practices and changes in innovation as startups grow, consistent with active learning models, prevalent in the literature. At the same time, we fail to find compelling evidence of conditional age effects, which would be consistent with an important role for passive learning models. As we discussed in the introduction, the absence of compelling evidence for passive learning models implies that financial intermediaries such as venture capital or private equity funds have the potential to substantially accelerate firm growth by providing financing and expertise to startups.

Third, we compare size effects of startups in manufacturing and services and find that non-monotonicities in firm size matter. These non-monotonicities are consistent with the view that organizational capital accumulation and innovation run into diminishing returns in firm size, which one might model as non-monotonic scale effects in static production functions and innovation production functions. A model based on these stylized facts has the potential to endogenously generate mean-reversion in firm performance as large firms are bound to run into these diminishing returns. The contrast of innovation size effects between manufacturing and services also has the potential to shed light in firm dynamics more generally. The difference between size effects of innovation in manufacturing vs services suggest that the present value of growth opportunities is quite different for manufacturing as compared to services firms. Furthermore, there are potentially important aggregate consequences of non-monotonic size effects of innovation, if economic activity shifts from manufacturing to services over time. We leave those questions for future research.

References

Acemoglu, D., P. Aghion, C. Lelarge, J. Van Reenen, and F. Zilibotti, "Technology, Information, and the Decentralization of the Firm," *Quarterly Journal of Economics*, 2007, *122* (4).

Aghion, P. and J. Tirole, "Formal and real authority in organizations," *Journal of Political Economy*, 1997.

__, U. Akcigit, and P. Howitt, "What do we learn from Schumpeterian Growth Theory," *Handbook of Economic Growth*, 2013.

Akcigit, U. and W. Kerr, "Growth through Heterogeneous Innovations," *NBER Working Paper*, 2010.

__, H. Alp, and M. Peters, "Lack of Selection and Imperfect Managerial Contracts: Firm Dynamics in Developing Countries," *mimeo, University of Pennsylvania*, 2014.

Atkeson, Andrew and Patrick Kehoe, "Modeling and Measuring Organization Capital," *Journal of Political Economy*, 2005.

Black, Sandra E. and Philip E. Strahan, "Entrepreneurship and Bank Credit Availability," *Journal of Finance*, 2002, *57* (6), 2807–2833.

Bloom, N., C. Jones, J. Van Reenen, and M. Webb, "Are ideas getting harder to find?," *Working Paper, Stanford University*, 2016.

— , L. Garicano, R. Sadun, and J. Van Reenen, "The distinct effects of information technology and communication technology on firm organization," *Management Science*, 2013.

Bresnahan, T., E. Brynjolfsson, and L. Hitt, "Information technology, workplace organization and the demand for skilled labor: firm-level evidence," *Quarterly Jounral of Economics*, 2002.

Caliendo, Lorenzo and Esteban Rossi-Hansberg, "The Impact of Trade on Organization and Productivity," *Quarterly Journal of Economics*, 2012, *1393*, 1467.

— , Ferdinando Monte, and Esteban Rossi-Hansberg, "The Anatomy of French Production Hierarchies," *Journal of Political Economy*, 2015, *123* (4), 809–852.

Calvo, Guillermo A and Stanislaw Wellisz, "Supervision, Loss of Control, and the Optimum Size of the Firm," *Journal of Political Economy*, 1978, pp. 943–952.

Deaton, Angus, *The Analysis of Household Surveys*, Johns Hopkins University Press, 1997.

Dunne, Timothy, Mark Roberts, and Larry Samuelson, "The Growth and Failure of U.S. Manufacturing Plants," *Quarterly Journal of Economics*, 1989.

Ericson, Richard and Ariel Pakes, "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies*, 1995, *62* (1), 53–82.

Evans, David S. and Boyan Jovanovic, "An Estimated Model of Entrepreneurial Choice Under Liquidity Constraints," *Journal of Political Economy*, 1989, pp. 808–827.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "The Slow Growth of New Plants:

Learning about Demand?," *mimeo, Chicago Booth*, 2013.

Garcia, A, J. Jaumandreu, and C. Rodriguez, "Innovation and jobs: evidence from manufacturing firms," *Working Paper, Boston University*, 2005.

Garetto, S., L. Oldenski, and N. Ramondo, "Life-Cycle Dynamics and the Expansion Strategies of US Multinational Firms," *Working Paper, Boston University*, 2016.

Garicano, L., "Hierachies and the organization of knowledge in production," *Journal of Political Economy*, 2000.

_ and E. Rossi-Hansberg, "Organizing Growth," *Journal of Economic Theory*, 2012.

Hall, Bronwyn, Jacques Mairesse, and Laura Turner, "Identifying Age, Cohort and Period Effects in Scientific Research Productivity: Discussion and Illustration Using Simulated and Actual Data on French Physicists," *Economics of Innovation and New Technology*, 2007.

Haltiwanger, John, Ron Jarmin, and Javier Miranda, "Who Creates Jobs? Small versus Large versus Young," *Review of Economics and Statistics*, 2013.

Hellmann, Thomas and Manju Puri, "Venture Capital and the Professionalization of Start-Up Firms: Empirical Evidence," *Journal of Finance*, 2002, *57* (1), 169–197.

Holmstrom, B., "Agency Costs and Innovation," *Journal of Economic Behavior and Organization*, 1989.

Holtz-Eakin, Douglas, David Joulfaian, and Harvey S Rosen, "Sticking it Out: Entrepreneurial Survival and Liquidity Constraints," *Journal of Political Economy*, 1994, pp. 53–75.

Hong, Bryan, Lorenz Kueng, and Mu-Jeung Yang, "Complementarity of Performance Pay and Task Allocation," *mimeo, University of Washington*, 2016.

Hsieh, C. and Peter J. Klenow, "The Life Cycle of Plants in India and Mexico," *mimeo Stanford University*, 2014.

Hurst, Erik and Annamaria Lusardi, "Liquidity Constraints, Household Wealth, and Entrepreneurship," *Journal of Political Economy*, 2004, *112* (2), 319–347.

Jovanovic, B., "Selection and the Evolution of Industry," *Econometrica*, 1982.

Klepper, S., "Entry, Exit, Growth and Innovation over the Product Life Cycle," *American Economic Review*, 1996.

Klette, T. and S. Kortum, "Innovating firms and aggregate innovation," *Journal of Political Economy*, 2004.

Mairesse, Jacques and Pierre Mohnen, "Using Innovations Surveys for Econometric Analysis," *NBER Working Paper*, 2010.

Manso, Gustavo, "Motivating Onnovation," *Journal of Finance*, 2011, *66* (5), 1823–1860.

OECD, Eurostat, and EC, "The measurement of scientific and technological activities," *https://www.oecd.org/sti/inno/2367580.pdf*, 1992.

Ouimet, Paige and Rebecca Zarutskie, "Who Works for Startups? The Relation between Firm Age, Employee Age, and Growth," *Journal of Financial Economics*, 2014, *112* (3), 386–407.

Peters, B., "Innovation and firm performance: An empirical investigation for German firms," *ZEW Economic Studies*, 2008.

Puri, Manju and Rebecca Zarutskie, "On the Life Cycle Dynamics of Venture-Capital- and Non-Venture-Capital-Financed Firms," *Journal of Finance*, 2012, *67* (6), 2247–2293.

Rajan, R. and L. Zingales, "The Firm as a Dedicated Hierarchy: A Theory of the Origins and Growth of Firms," *Quarterly Journal of Economics*, 2001, *116*, 805–851.

Reenen, J. Van, "Employment and technological innovation: evidence from UK manufacturing firms," *Journal of Labour Economics*, 1997.

Schulhofer-Wohl, Sam, "The Age-Time-Cohort Problem and the Identification of Structural Parameters in Life-Cycle Models," *mimeo, Federal Reserve Bank of Minneapolis*, 2013.

Sutton, John, "Gibrat's Legacy," *Journal of Economic Literature*, 1997.

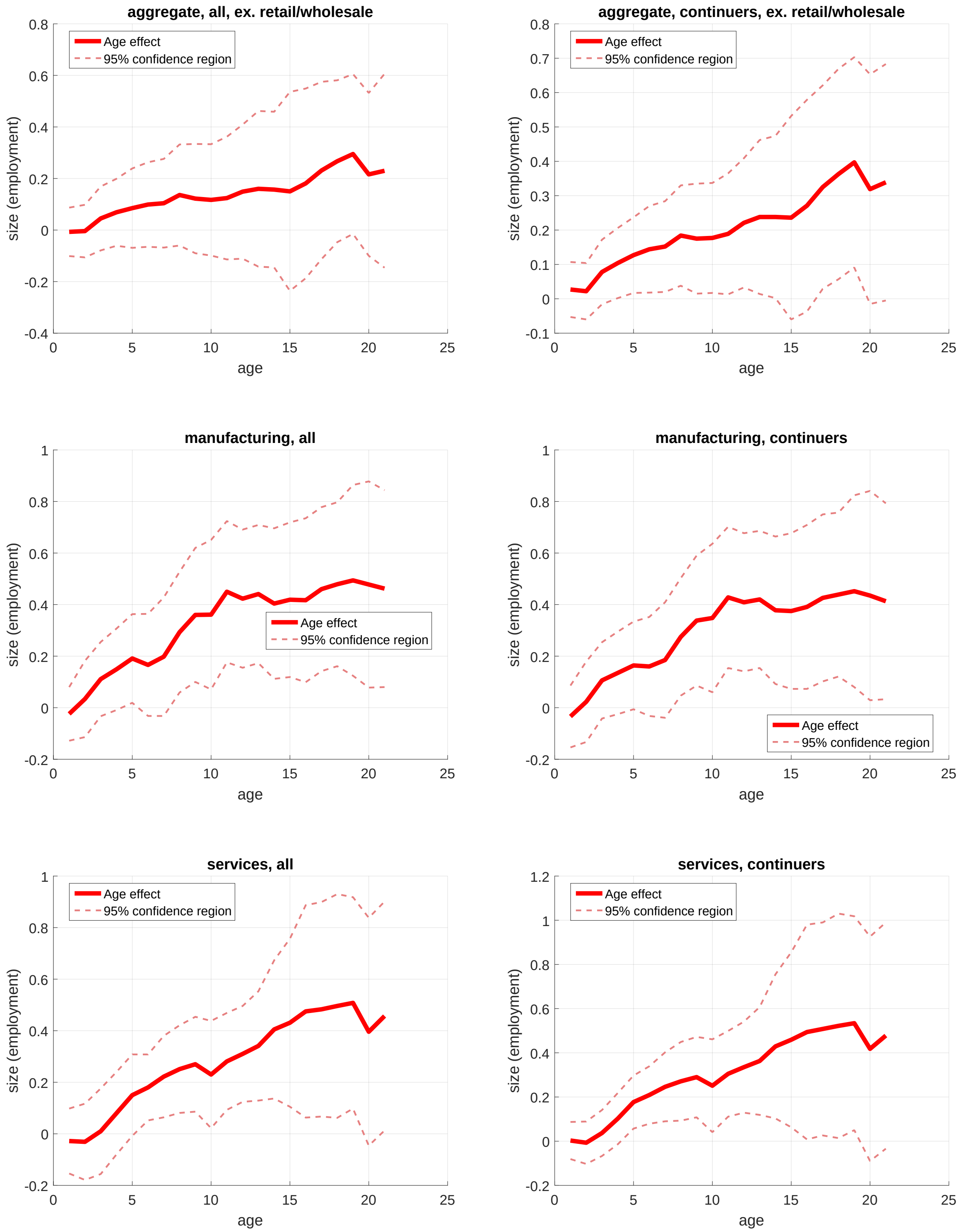

Figure 1 – Estimates are from a regression of log number of employees on dummies of age, size and establishment fixed effects: $\ln(s_{e,\{a,c,t\}}) = \nu + \sum_{a=1}^{A} \lambda_a \cdot D_a + \sum_{t=1}^{T} \tau_t \cdot D_t + \sum_{e=1}^{E_t} \xi_e \cdot D_e + \texttt{error}$. Standard errors are clustered at the size, region, industry level, corresponding to sample stratification of population weights. Subsamples are defined in graph titles.

Figure 2 – Estimates are from a regression of log number of employees on dummies of age, size and establishment fixed effects: $\ln(s_{e,\{a,c,t\}}) = \nu + \sum_{a=1}^{A} \lambda_a \cdot D_a + \sum_{t=1}^{T} \tau_t \cdot D_t + \sum_{e=1}^{E_t} \xi_e \cdot D_e + \texttt{error}$. Standard errors are clustered at the size, region, industry level, corresponding to sample stratification of population weights. Subsamples are defined in graph titles.

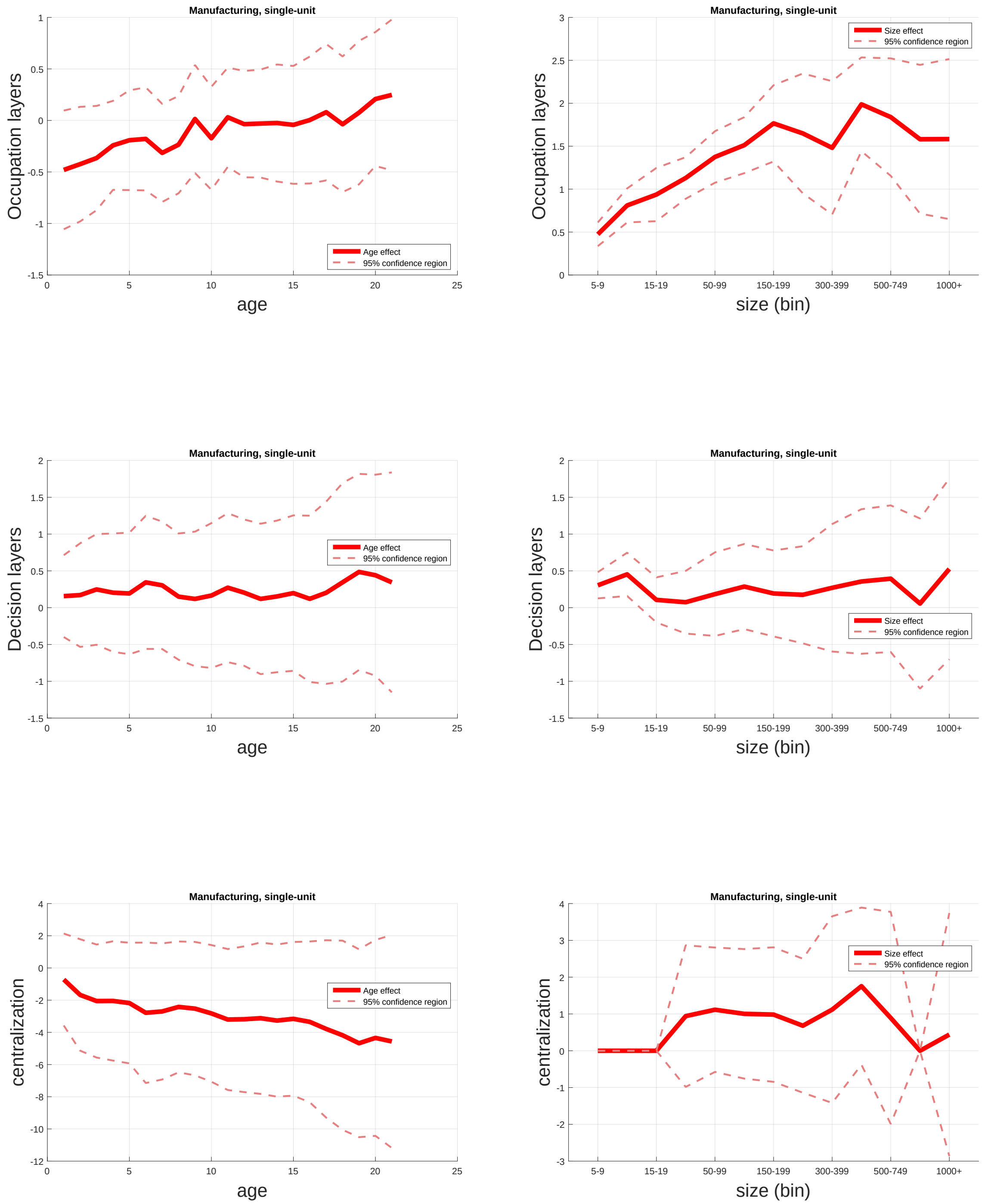

Figure 3 – Estimates are from a regression of management practices on dummies of age, size and establishment fixed effects: $\ln(x_{e,\{a,c,t,s\}}) = \nu + \sum_{s=1}^{S} \kappa_s \cdot D_s + \sum_{a=1}^{A} \lambda_a \cdot D_a + \sum_{t=1}^{T} \tau_t \cdot D_t + \sum_{e=1}^{E_t} \xi_e \cdot D_e + \texttt{error}$. Dependent variables $x_{e,\{a,c,t,s\}}$ are displayed on the y-axis, see main text for precise definitions. Standard errors are clustered at the size, region, industry level, corresponding to sample stratification of population weights. Subsamples are defined in graph titles. Single-unit, manufacturing firms: age (left) vs. size (right) effects.

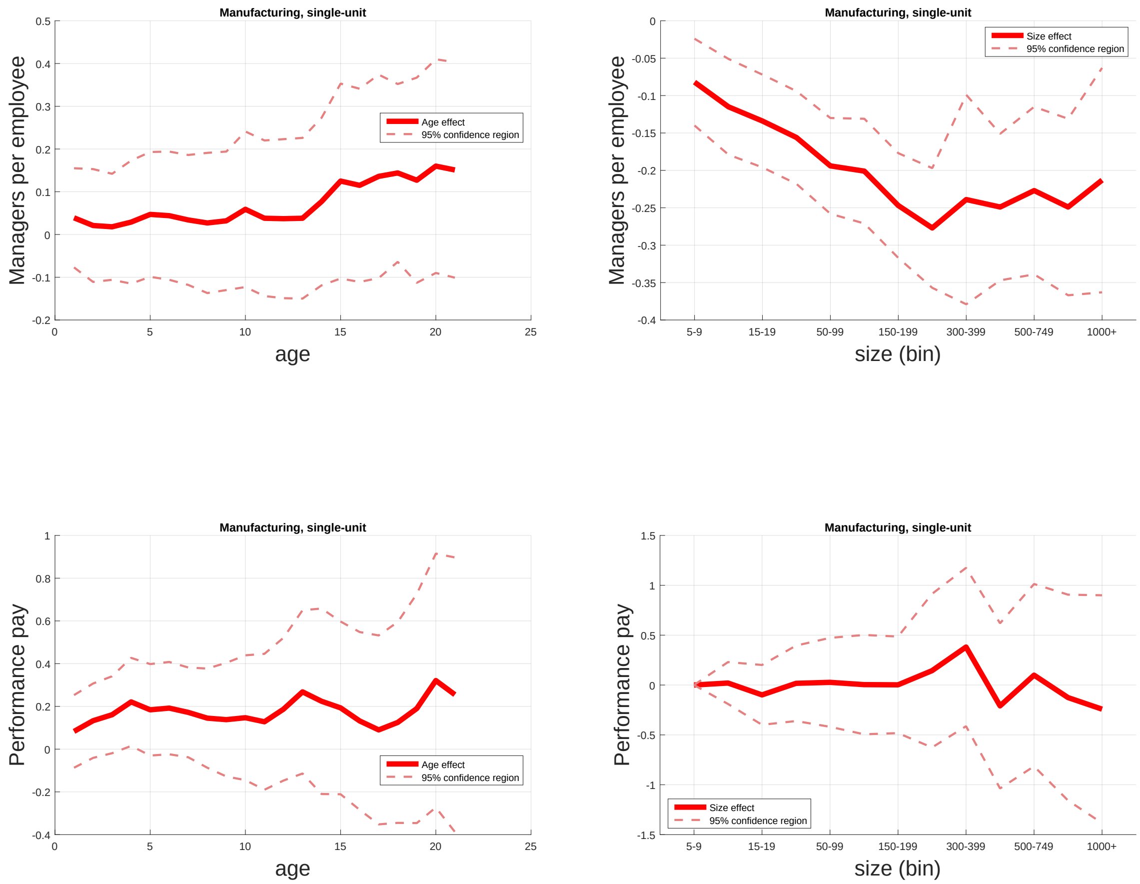

Figure 4 – Estimates are from a regression of management practices on dummies of age, size and establishment fixed effects: $\ln(x_{e,\{a,c,t,s\}}) = \nu + \sum_{s=1}^{S} \kappa_s \cdot D_s + \sum_{a=1}^{A} \lambda_a \cdot D_a + \sum_{t=1}^{T} \tau_t \cdot D_t + \sum_{e=1}^{E_t} \xi_e \cdot D_e + \texttt{error}$. Dependent variables $x_{e,\{a,c,t,s\}}$ are displayed on the y-axis, see main text for precise definitions. Standard errors are clustered at the size, region, industry level, corresponding to sample stratification of population weights. Subsamples are defined in graph titles. Single-unit, manufacturing firms: age (left) vs. size (right) effects.

Figure 5 – Estimates are from a regression innovation measures on dummies of age, size and establishment fixed effects: $\ln(x_{e,\{a,c,t,s\}}) = \nu + \sum_{s=1}^{S} \kappa_s \cdot D_s + \sum_{a=1}^{A} \lambda_a \cdot D_a + \sum_{t=1}^{T} \tau_t \cdot D_t + \sum_{e=1}^{E_t} \xi_e \cdot D_e + \texttt{error}$. Dependent variables $x_{e,\{a,c,t,s\}}$ are displayed on the y-axis, see main text for precise definitions. Standard errors are clustered at the size, region, industry level, corresponding to sample stratification of population weights. Subsamples are defined in graph titles. Single-unit, manufacturing firms: age (left) vs. size (right) effects.

Figure 6 – Estimates are from a regression of management practices on dummies of age, size and establishment fixed effects: $\ln(x_{e,\{a,c,t,s\}}) = \nu + \sum_{s=1}^{S} \kappa_s \cdot D_s + \sum_{a=1}^{A} \lambda_a \cdot D_a + \sum_{t=1}^{T} \tau_t \cdot D_t + \sum_{e=1}^{E_t} \xi_e \cdot D_e + \texttt{error}$. Dependent variables $x_{e,\{a,c,t,s\}}$ are displayed on the y-axis, see main text for precise definitions. Standard errors are clustered at the size, region, industry level, corresponding to sample stratification of population weights. Subsamples are defined in graph titles. Single-unit firms: mfg (left) vs. serv (right) industries

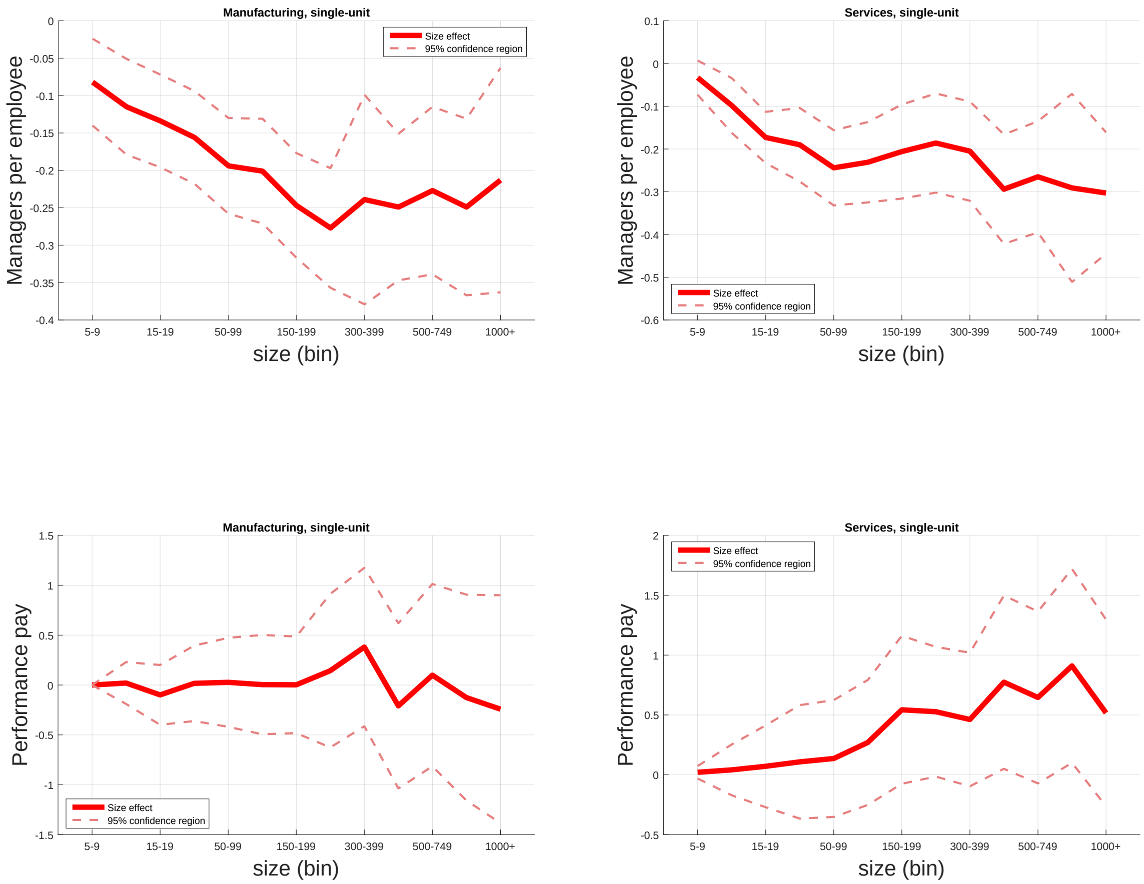

Figure 7 – Estimates are from a regression of management practices on dummies of age, size and establishment fixed effects: $\ln(x_{e,\{a,c,t,s\}}) = \nu + \sum_{s=1}^{S} \kappa_s \cdot D_s + \sum_{a=1}^{A} \lambda_a \cdot D_a + \sum_{t=1}^{T} \tau_t \cdot D_t + \sum_{e=1}^{E_t} \xi_e \cdot D_e + \texttt{error}$. Dependent variables $x_{e,\{a,c,t,s\}}$ are displayed on the y-axis, see main text for precise definitions. Standard errors are clustered at the size, region, industry level, corresponding to sample stratification of population weights. Subsamples are defined in graph titles. Single-unit firms: mfg (left) vs. serv (right) industries

Figure 8 – Estimates are from a regression of innovation measures on dummies of age, size and establishment fixed effects: $\ln(x_{e,\{a,c,t,s\}}) = \nu + \sum_{s=1}^{S} \kappa_s \cdot D_s + \sum_{a=1}^{A} \lambda_a \cdot D_a + \sum_{t=1}^{T} \tau_t \cdot D_t + \sum_{e=1}^{E_t} \xi_e \cdot D_e + \texttt{error}$. Dependent variables $x_{e,\{a,c,t,s\}}$ are displayed on the y-axis, see main text for precise definitions. Standard errors are clustered at the size, region, industry level, corresponding to sample stratification of population weights. Subsamples are defined in graph titles. Single-unit firms: mfg (left) vs. serv (right) industries