A Tale of Two C(...)s:
Competence and Complementarity

Simeon D. Alder[*]   Fane N. Groes[†]
University of Notre Dame   Copenhagen Business School

October 13, 2014

PRELIMINARY AND INCOMPLETE – PLEASE DO NOT CITE

Abstract

We build a tractable assignment model to characterize the matching and separation patterns of CEOs and their employers. Managers learn about their own type by observing a sequence of public signals. Each period, the sorting of firms and managers is *ex ante* perfectly assortative, but is generally not so *ex post*. We calibrate the model to match empirical targets from a large matched employer-employee dataset covering the Danish labor force between 2000 and 2009. We exploit the non-monotonicity of executive compensation in the employer type to parameterize the model, among them the degree of complementarity between the characteristics of the manager and those associated with the firm in the production function. The results fill an important gap in the literature on the aggregate effects of mismatch and our theory is a natural building block toward a dynamic theory of entrepreneurship.

JEL Codes: C78, D33, E23, J31, O47

Keywords: Sorting, complementarity, occupational choice, identification, firm size distribution, aggregate productivity

---
[*]Email: simeon.alder.1@nd.edu
[†]Email: fg.eco@cbs.dk

1 Introduction

A sizeable literature addresses the allocation of skills across firms or jobs. Beginning with Lucas (1978) it has been argued that differences in technologies or organization capital, for instance, are key to understanding the size distribution of firms and establishments. Similarly, individuals are endowed with different innate abilities, which translate into differences in market-relevant skills through education or work experience, to name just a few. While heterogeneity in productivities and skills is an uncontroversial idea, the extent to which complementarities govern their interaction remains an open question. Abowd et al. (1999) find no empirical evidence for worker-job complementarities. Lopes de Melo (2013), on the other hand, argues that the estimates from fixed effects regressions may be biased due to the non-monotonicity of wages in firm types, a feature that Eeckhout and Kircher (2011) also emphasize. Lise et al. (2013) find evidence for positive complementarity, at least for relatively well educated workers in the United States.

But why do we care about complementarities at all? With complementarities, a worker's marginal product depends on her own skill and on the characteristics of the job – say, the task-specific technology in use – she is assigned to. Clearly then, the optimal allocation of workers to jobs and the prices that decentralize these assignments critically depend on the degree to which skills and technologies complement each other. What's more, achieving aggregate efficiency is not just a matter of making sure everyone has *a* job but requires that everybody ends up with the *right* kind of job.

The intuition is quite simple and to fix ideas consider the problem of assigning a manager to a firm. In the absence of complementarities, a CEO's marginal product is a function of the efficiency units she embodies only. Her contribution to total surplus is the same across firms and the assignment of managerial talent is of no consequence for efficiency. The size of the economy is governed by the aggregate supply of talent. In contrast, if managers and firms are complements, a competent CEO's marginal product rises more sharply across different productivity levels than that of an incompetent manager and assortative matching maximizes the mutual leverage of marginal products. In this environment, individual (or marginal) effects are characterized by a combination of the primitives (type distributions and technology) and the equilibrium itself (assignment). The challenge is to write down a model that enables us to disentangle the effects driven by complementarities from those that depend on the distribution of types. This is a worthwhile endeavor since we have reason to believe that complementarities matter.

Consider, for example, a Canadian street performer by the name of Guy Laliberté. He started his career as an accordeon player, fire breather, and stiltwalker with a small group of colorful

characters in the small Canadian town of Baie-Saint-Paul. In due time he founded *Cirque du Soleil*, which is now a global enterprise with several thousand employees. While Guy reportedly was extraordinarily creative from an early age, his marginal product rose sharply over the course of his career. Arguably, it was low while he was "assigned" to the sidewalks of Baie-Saint-Paul. He quit the sidewalks and moved into the first *Cirque du Soleil* tent seating 800 spectators in 1984. After several upgrades, the circus is now on multiple simultaneous tours in addition to resident shows in Las Vegas, Los Angeles, and Orlando. His creativity is now matched with a more productive technology that allows him to entertain millions of spectators rather than the few hundred or thousand pedestrians who saw him on a street corner in his early days. By 2009, his marginal product was high enough to buy him a $35 million ticket to the *International Space Station*.

In the absence of talent-technology complementarities, Guy's marginal product is proportional to his skills and that, in turn, would suggest that his performance talent increased by a few orders of magnitude over the course of three decades. While we acknowledge the importance of endogenous skill accumulation, we argue that complementarities are a critical feature of talent markets in general and the market for entrepreneurs and senior managers in particular.

If CEO and firm types were observable, the model parameterization would be straightforward. Alas, they are not and previous assignment models have struggled to separately identify the distributions of types from the properties of the technology that generates the match surplus for precisely this reason (Gabaix and Landier, 2008; Terviö, 2008; Alder, 2013). The amount of identification we can squeeze out of the cross-section of prices, namely the compensation of CEOs and payments to the owners of the projects/firms, is limited.

To make further progress we need a dynamic assignment model with endogenous separation and turnover to harness information that is embedded in the time series dimension of the data. To learn something about Guy Laliberté's creative talent separately from the contribution of the technology he operates, we need to know the 1980s busker as well as the successful CEO from the 1990s and 2000s. Our theory accomplishes this goal in a tractable way and we parameterize the model using a matched employer-employee data from three different sources that cover the Danish labor force from 2000 to 2009. The register data contains socioeconomic variables, employment information, and a complete employer-employee link for the entire population between 1999 and 2009. The *Købmandstadens Oplysningsbureau* (KOB) dataset contains firm accounting data for all Danish limited liability firms from 1991 to the present. Lastly, the *Erhvervs- og Selskabstyrelsen* (ES) data identifies CEOs and board members of limited liability firms from 2000 to 2010. Unique firm and worker identifiers enable us to merge information from these three sources into a single dataset and to construct

employment histories for virtually all individuals of interest to us. This rich panel structure delivers the additional identification we need to jointly estimate the parameters of interest in the model.

In our model agents do not know their own type but receive a sequence of public signals that enables them to update their beliefs using Bayes' rule (as in Groes et al., 2012). In contrast to Jovanovic (1979) agents learn about their own abilities rather than about an *ex ante* unknown match quality.[1] While the equilibrium features perfect assortative matching in the agents' *priors*, it does not do so *ex post*. Separation and matching are friction- and costless. The individuals' true types are distributed normally and the firm characteristics are drawn from a discrete distribution. CEOs separate whenever the gap between priors and posteriors is big enough to warrant a switch to a different project type. Separation and matching are governed by the price system and the CEOs' decisions are self-selective. Each period, participation in the market for executive talent is endogenous and CEOs pick the employer that offers the most lucrative wage contract in expectation. The model does not distinguish between quits and layoffs. This being said, *ex post* realizations that are associated with low CEO pay have a natural interpretation as layoffs whereas payouts above a certain threshold trigger separations that have the look and feel of quits. As in Eeckhout and Kircher (2011) and Groes et al. (2012) compensation is non-monotonic in the project type: the equilibrium wage offers are such that for each CEO – of a certain age and with some belief about her own abilities – there exists a project that offers a more lucrative contract than any other. Importantly, the ideal type is not typically the most productive project. We exploit this non-monotonicity result to discipline the model's sorting and separation patterns using a panel dataset containing virtually all Danish CEOs and their employers.

Unlike Eeckhout and Kircher (2011) we use wage (i.e. CEO compensation) data as well as data on profits and output. While they argue that firm performance does not respond much to variations in the productivity of individual workers inside the firm, Bennedsen et al. (2008, 2012) find that CEOs matter. They have a measurable impact on standard measures of firm performance. Moreover, CEOs are unique in this respect: other senior executives do not significantly affect performance. This empirical evidence supports our model's implicit assumption that, in principle, the CEO's contribution (or lack thereof) can impact firm performance. By using payment data on both sides of the market we can characterize the strength *and* the sign of the equilibrium assignment. What's more, our model qualitatively matches the stylized fact that firm performance is an important determinant of hiring, retention, and separation decisions in the market for chief executives and, possibly, entrepreneurs.[2] We use a simulated method of moments to parameterize the model for the best fit with our data.

---

[1]See also Nagypal (2007) for a similar distinction.
[2]See, for instance, the discussion in Eisfeldt and Kuhnen (2013).

Our theory has potential applications beyond the CEO-to-project assignments considered here. In particular, it is a natural building block toward a dynamic theory of entrepreneurship.

In our model, the returns to entrepreneurship are uncertain while the outside option is associated with a constant flow payment.[3] Occupational choice theories like Johnson (1978), Jovanovic (1979), or Miller (1984) suggest that individuals select risky occupations when they are young. Evans and Jovanovic (1989), in contrast, argue that liquidity constraints are tighter for younger individuals since they have, on average, accumulated less wealth. This constraint acts as a countervailing force to the rookies' appetite for risk. Evans and Leighton (1989) find that the hazard rate into self-employment is constant in age and it is natural to conclude that liquidity constraints are binding in the data. Our theory offers alternative support for the Knightian view according to which risk-bearing is a defining trait of entrepreneurship (Knight, 1921). Rather than liquidity constraints, our theory's countervailing force is fueled by the fact that inexperienced entrepreneurs have less precise beliefs about their own abilities. With diminishing marginal returns to talent the market is more selective for young entrepreneurs compared to older ones.

We emphasize that our model is a building block rather than a full-fledged theory of entrepreneurship. Most importantly, we abstract from entry, exit, and other project dynamics that are distinct from those of the entrepreneurs themselves. We believe, however, that this is a promising direction for future research.

The remainder of the paper is organized as follows. Section 2 details the model environment and defines the equilibrium. Section 3 characterizes the patterns of CEO mobility. We describe the data in section 4 and the specifics of our structural estimation in section 5. The counterfactuals in section 6 quantify the effects of mismatch. Section 7 concludes.

2 Model

2.1 Population and Endowment

The economy is populated by a unit measure of individuals who are endowed (at birth) with skill a. This ability to manage a firm is drawn from a known normal distribution with mean zero and precision ϕ. The draws from this distribution are i.i.d. across agents.

Individuals live for S periods. Each period, a new cohort of size $\frac{1}{S}$ enters the economy, while the oldest cohort (of the same size) retires. New entrants do not know their type. Instead, they observe a *public* signal $a_0 = a + \alpha_0$. The innovation α_0 is drawn from a normal distribution

---

[3]One may think of the outside option as the wage associated with supplying a raw unit of labor, as in Lucas' canonical span-of-control model (Lucas, 1978).

with mean zero and precision ψ. At time $t \in \{1, 2, \ldots\}$, a manager's contemporaneous ability is given by $a_t = a + \alpha_t$. The uncertainty associated with the CEOs' type, α_t, is drawn from the same distribution as α_0.[4] The innovations are independent cross-sectionally and over time. In addition, individuals have a periodic unit endowment of time and each period they decide what fraction to spend on managerial tasks and on an alternative occupation or activity that is compensated at the exogenous rate w per unit of time.

The economy is also endowed with a unit measure of long-lived projects q, drawn from a discrete c.d.f. $F(\cdot)$ with K distinct values and P.M.F. $f(q_k) = \gamma_k$, for $k \in \{1, \ldots, K\}$. The corresponding C.D.F. is $F(q_k) = \sum_{\ell=1}^{k} \gamma_\ell = \Gamma_k$. Each individual owns one such project for the duration of her lifetime. When generation s dies after T periods, the "orphaned" projects are bequeathed to randomly chosen individuals of the cohort with birth date $s + T$.

2.2 Preferences and Technology

Individuals have linear preferences over their lifetime consumption stream:

$$U\big(\{c_t\}\big) = \sum_{s=t}^{t+T} \beta^{s-t} c_s$$

Since there is no consumption-saving tradeoff, c_s is simply capped by the sum of labor and managerial income as well as dividend payments and capital gains from firm ownership. A CEO-firm pair produces a contemporaneous surplus x_t, which is a function of their respective contributions a_t and q:

$$x_t = x(a + \alpha_t, q) \tag{1}$$

The technology is continuous and increasing in both the manager's effective contribution a_t and the project's attribute q. Moreover, $x(\cdot, \cdot)$ satisfies increasing differences.

2.3 Beliefs

The distribution of true, but unobserved managerial talent is stationary. Each cohort of size $\frac{1}{S}$ draws types from the same normal distribution with mean zero and variance $\frac{1}{\phi}$. Clearly then, the aggregate distribution of talent follows that same distribution. In contrast, individual realizations of a are not observable. Instead, each agent receives a public signal $a_t = a + \alpha_t$ and forms a belief about her own type. Since a_t is public information, everybody's beliefs about any one individual coincide. Recall that α_t is drawn from a mean-zero normal distribution

---

[4]Relaxing this assumption generates the same qualitative results but requires heavier notation.

with precision ψ:

$$\alpha_t \sim \mathrm{N}\left(0, \tfrac{1}{\psi}\right), \text{ for all } t \tag{2}$$

Each signal enables individuals to update their prior using a Kalman filter. Since the innovation is normal, the *posterior* of someone who is s years old – denoted by $\tilde{a}_{s+1}$ – is normal with mean

$$\hat{a}(s+1) \equiv \mathbb{E}\left(\tilde{a}(s+1)\right) = \tfrac{\phi_i + s\psi}{\phi_i + (s+1)\psi}\hat{a}(s) + \tfrac{\psi}{\phi_i + (s+1)\psi}a_t \tag{3}$$

and variance

$$\mathrm{Var}\left(\tilde{a}(s+1)\right) = \tfrac{1}{\phi + (s+1)\psi} \tag{4}$$

Note that the variance of the posterior is independent of the individuals life-time history of shocks and decreasing in age: older workers have a more precise idea about their own abilities. Since there is no private information, this implies that everybody in this economy has more precise beliefs about the abilities of older workers compared to younger ones. Since $\frac{\phi + s\psi}{\phi + (s+1)\psi} + \frac{\psi}{\phi + (s+1)\psi} = 1$ and the second term is decreasing in s, agents put more and more weight on their prior $\hat{a}(s)$ at the expense of the innovation a_t as they approach retirement at age S.

2.4 Wage Contracts

As in Groes et al. (2012), we consider the set of contracts that is bounded by two polar types, where either the owner or the manager assumes all risks associated with a particular match:

1. The wage offers are output-contingent contracts. The firm holds on to a reservation profit and the CEO is the residual claimant (and hence assumes all the risk).

2. The wage offers are *ex ante*-type-contingent contracts. The firm offers a non-contingent payment to the CEO and claims the residual surplus.

In the former case, firms determine the reservation profit that extracts the maximum amount of surplus subject to market clearing for prospective CEOs. The K different reservation profits only rely on the cross-sectional distribution of beliefs – which are common knowledge in this economy – and firm owners hire anyone who applies for the job. The assignment of CEOs is self-selective. In contrast, wage contracts that do not depend on the *ex post* surplus require that project owners know as much about the applicants as the CEOs themselves do. If they did not, CEOs would have incentives to misrepresent their type in order to extract a bigger share of the match value. Both contracts are efficient in that they decentralize the planner's solution that features *ex ante* assortative matching.

The realized split of the surplus between managerial compensation and payments to share-holders depends on the type of contract. In particular, the joint pattern of pay volatility and switching (turnover) differs between the two types of contracts. When pay is output-contingent then deviations from expected compensation are associated with an increased probability of separation. In contrast, turnover is orthogonal to contemporaneous pay, but correlated with firm profits when the terms of the contract are type-contingent. The data we have suggests that, in practice, the wage contract is somewhere between these types. We begin by spelling out the wage offers in the polar self-selection and full insurance cases.

Recall that beliefs about a CEO's type are denoted by $\tilde{a}(\hat{a}, s)$ and have two elements: (1) the expectation $\hat{a}$, which is a discrete-time Martingale, and (2) the variance associated with a CEO of a certain age, say s. The variance $\frac{1}{\phi+s\psi}$ decreases deterministically as s goes from zero to S. In the special case of our model where $x(a_t, \cdot)$ is linear in a_t we can ignore the variance altogether and CEOs are sufficiently characterized by $\hat{a}$ for sorting purposes. When the technology exhibits strictly increasing differences, however, we need to keep track of a CEO's expectation *and* variance.[5] Since the innovation α is normal, the prior $\tilde{a}$ is also normal.[6]

$$\tilde{a}(\hat{a}, s) \sim \mathrm{N}\big(\hat{a}, \tfrac{1}{\phi+s\psi}\big) \tag{5}$$

In a slight abuse of notation, we denote the expected output of a manager with prior $\tilde{a}(\hat{a}, s)$ who is paired with project q by

$$\mathbb{E}\big[x\big(\tilde{a}(\hat{a}, s), q\big)\big] \equiv \int_{-\infty}^{\infty} x\big(a_t, q\big)\mathrm{d}F_{\hat{a},s}(a_t) \tag{6}$$

where $F_{\hat{a},s}$ is the C.D.F. of $\tilde{a}(\hat{a}, s)$.

2.4.1 Output-Contingent Offers

When the CEO's pay is contingent on output, she is a residual claimant and bears all the risk. Firms keep a reservation profit $\pi(q)$ for themselves and offer $\omega(x) = x(a + \alpha_t, q) - \pi(q)$ to prospective CEOs. Keep in mind that for particularly unfavorable realizations of $a + \alpha_t$, the CEO may earn $\omega(x) < 0$. When the firm retains $\pi(q)$, the sorting of CEOs relies on self-selection and prospective employers do not care about the applicant's type. While project owners want to maximize their share of the surplus, they must offer prospective CEOs terms

---

[5]Curvature requires us to project the expectation and variance of a manager's type onto a one-dimensional space in order to characterize an *ex ante* assortative match.

[6]Put differently, when the technology is super-modular we need to keep track of the entire distribution of $\tilde{a}(\hat{a}, s)$. Since it is normal, mean and variance describe the distribution completely. When the technology is linear, the expectation $\hat{a}$ is a sufficient statistic.

such that the vacancy is filled. That is, the K different $\{\pi_k\}_{k=1}^K$ must be such that the market for CEOs clears.

To ensure market clearing we need to verify that CEOs self-select exactly one employer, except for a finite number – $K \times T$ of them, to be precise – with measure zero. We call them "critical" types. To make further progress we need to describe the cross-sectional distribution of beliefs in the economy. To build some intuition, consider a cohort of age s. According to equation (5), the variance of the CEOs' beliefs *within* this cohort is constant and equal to $\frac{1}{\phi+s\psi}$. Since agents will differ in their expected ability $\hat{a}$, we need to characterize the cohort-specific distribution of these mean beliefs. One can show that mean beliefs of CEOs of age s are cross-sectionally normal:

$$\hat{a}(s) \sim \text{N}(0, \tfrac{s\psi}{\phi(\phi+s\psi)}) \tag{7}$$

We denote the corresponding C.D.F. by F_s. Recall that the true (but unobserved) distribution of abilities is $N\left(0, \frac{1}{\phi}\right)$. The cross-section of beliefs must be unbiased and therefore the distribution "inherits" the zero mean property. To understand the variance, let us compare the case of a CEO who has not yet observed a signal ($s = 0$) with one who has seen many of them ($s \to \infty$). A group of CEOs who have not received any signals about their type must have the same belief: zero. The distribution collapses to a mass point and the variance is indeed $\frac{0\psi}{\phi(\phi+0\psi)} = 0$. In contrast, the Bayesian update in equations (3) and (4) implies that the distribution of expected abilities converges to the distribution of true types a with a variance around each point estimate that decays to zero as $s \to \infty$. In the limit, CEOs know their own type for sure. The cross-sectional distribution of expected abilities approaches the variance of true types in the (very) long run:

$$\frac{s\psi}{\phi(\phi+s\psi)} \xrightarrow{s\to\infty} \frac{1}{\phi}$$

Since the sorting is *ex ante* assortative in our model, we find it helpful to rank CEOs by their expected ability $\hat{a}$ within their cohort. Moreover, it turns out to be useful to define the inverse C.D.F. of expected abilities in cohort s:

$$\hat{a}(s)[i] = F_s^{-1}(i) \tag{8}$$

The inverse C.D.F. maps a CEO's rank i and age s into an expected ability $\hat{a}$. When she is paired with a firm of type q_k, the match generates a value $\mathbb{E}\big[x\big(a_t(s,i), q_k\big)\big]$, where the expectation is over the normal realizations of a_t with mean $\hat{a}$ and variance $\frac{1}{\phi+s\psi}$. For a given K-tuple of reservation profits $\{\pi_k\}_{k=1}^K$ and age s, there are K critical ranks $\{i_k(s)\}_{k=1}^K$ and $i_0(s) \equiv 0$ (for all s) that satisfy:

$$\mathbb{E}\big[x\big(a_t(s, i_k(s)), q_{k-1}\big)\big] - \pi_{k-1} = \mathbb{E}\big[x\big(a_t(s, i_k(s)), q_k\big)\big] - \pi_k \qquad (9)$$

The "critical" CEO $i_k(s)$ is indifferent between the contracts offered by type q_k and q_{k-1} firms and $\frac{i_k(s) - i_{k-1}(s)}{S}$ denotes the measure of CEOs who select the contract offered by q_k. Since $x(\cdot, \cdot)$ exhibits increasing differences and given the beliefs about their own type, alternative employers q_{-k} offer less lucrative contracts in expectation. Summing the difference of "adjacent" critical ranks over all cohorts $s \in \{1, \ldots, S\}$ yields the total measure of CEOs who select the contract offered by q_k:

$$j_k - j_{k-1} = \tfrac{1}{S} \sum_{s=1}^{S} \Big(j_k(s) - j_{k-1}(s) \Big) \qquad (10)$$

In addition to the indifference condition, the contracts must satisfy a participation constraint and market clearing conditions. CEOs have an outside option that delivers w for sure.[7] Clearly, then, the marginal CEOs indexed by $j_{k^*}(s)$ and the marginal project q_{k^*} satisfy:

$$\mathbb{E}\big[x\big(a_t(s, i_k(s)), q_{k^*}\big)\big] = w + \pi_{k^*} \qquad (11)$$

Knowing the distribution of beliefs across all cohorts, firms choose the reservation profits $\{\pi_\ell\}_{\ell=k^*}^{K}$ that satisfy (11) and clear the market for CEOs:

$$j_{k^*+1} - j_{k^*} \leq \gamma_{k^*} \qquad (12)$$
$$j_{\ell+1} - j_\ell = \gamma_\ell, \text{ for } \ell = k^* + 1, \ldots, K - 1 \text{ and } j_{K+1} = 1 \qquad (13)$$

We assume that there is a sufficiently large supply of q_{k^*} so that the inequality in (12) is strict. In this case, q_{k^*} firms that sit idle – and hence produce no surplus or profits – bid down π_{k^*} and draw additional CEOs into the market until both the marginal firm and CEO are paid their outside options, 0 and w, respectively. Without loss of generality, we can relabel the projects by assuming that $k^* = 1$ and now K denotes the number of firm types that actually participate in the assignment market.

Definition 1 (Equilibrium with Output-Contingency) *An equilibrium is a set of reservation profits $\{\pi_k\}_{k=1}^{K}$ that satisfy the participation constraint in (11) and the market clearing conditions in equations (12) and (13) by way of the indifference conditions in (9) and the cross-cohort aggregation*

---

[7]The equilibrium is partial. It is straightforward to characterize the general equilibrium in a version of the theory with a Lucas (1978) span-of-control production function $x(\cdot, \cdot)^{1-\gamma} \ell^\gamma$ that requires managers as well as workers to produce the final good. In such an economy workers earn a competitive wage w, which is determined in general equilibrium.

in (10).

Since agents are risk-neutral, they select projects based on expected ability and compensation. Figures 1 and 2 plot the expected output-contingent compensation against the expected contribution against the CEO's belief about his own type for a simple version of the model with $K = 3$, $T = 5$ and outside option $w = 0.986$. This competitive wage implies that $i_1 = 0.95$ and only the right tail of the distribution is sufficiently competent to manage a project. Importantly, after controlling for expected ability, older CEOs receive higher compensation than young ones. Since they have more precise beliefs about their own type, they can command higher compensation in expectation when the production technology exhibits diminishing marginal returns to ability. As $\rho \to \infty$, precision ceases to matter and the expected compensation is determined by the expected ability alone.

Figures 3 and 4 highlight the *ex ante* selection: CEOs never accept a job that lies inside the envelope covering the K different contracts. The lower panel shows that older cohorts have more mass in the tail of the belief distribution and the occupational cutoff for the youngest CEOs is higher (1.77) than for the oldest (1.42). Once the uncertainty is realized, of course, the output-contingent compensation need not lie on the envelope. The plot of realized ability $a + \alpha_t$ against actual compensation lies *on* the wage offer curve, but not necessarily on the *envelope* of all offers. Clearly, CEOs bear all the risk. Alternatively, firms may offer a type-contingent contract.

2.4.2 Type-Contingent Offers

A CEO *type* is characterized by the pair (s, i), that is, an age s (associated with belief precision $\phi + s\psi$) and a rank i (corresponding to expected ability $\hat{a}(s)[i]$). When the offers are type-contingent, the firms are residual claimants and assume all the risk. Instead of picking the reservation profits $\{\pi_k\}_{k=1}^{K}$, firms offer a set of wage functions $\{\omega_k(s, i)\}_{k=1}^{K}$. Assuming that firm k knows as much about prospective CEOs as the candidates themselves do, the offers are such that it is indifferent between all of them *ex ante*:

$$\mathbb{E}\big[x\big(a_t(s, i), q_k\big)\big] - \omega_k(s, i) = \mathbb{E}\big[x\big(a_t(s', i'), q_k\big)\big] - \omega_k(s', i') \tag{14}$$

for $(i, i') \in [0, 1]$, $(s, s') \in \{1, \dots, S\}$, and $k \in \{1, \dots, K\}$.

CEOs, on the other hand, are not indifferent between prospective employers except, again, for $K \times S$ of them with measure zero. These critical CEOs, indexed by $i_k(s)$ as in section 2.4.1, are indifferent between the contracts offered by firm k and $k - 1$:

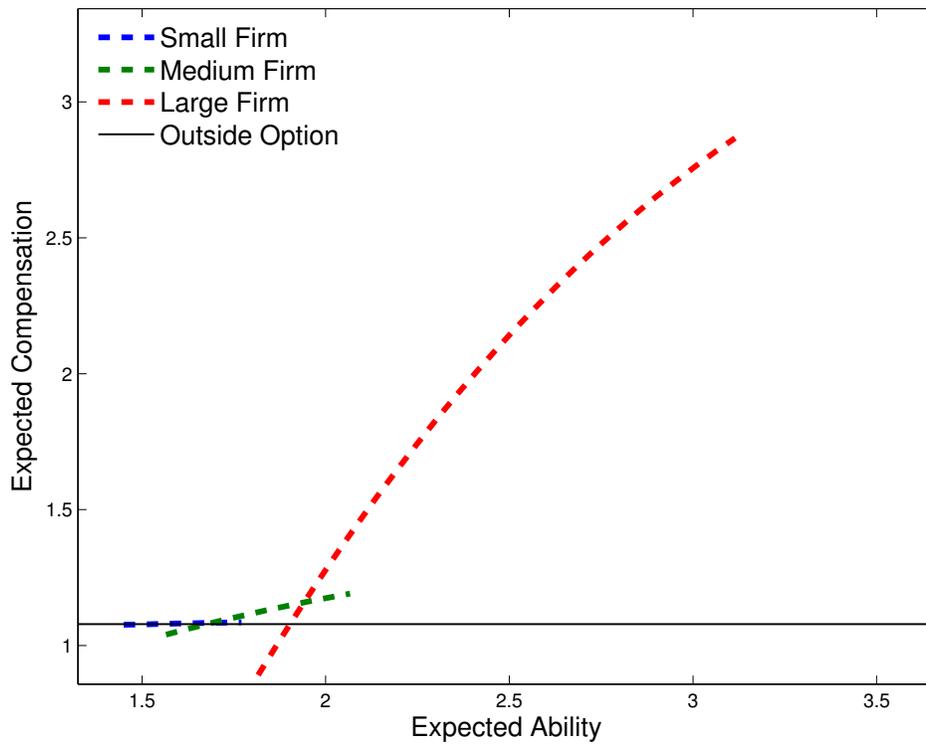$$\omega_k\big(s, i_k(s)\big) = \omega_{k-1}\big(s, i_k(s)\big) \tag{15}$$

11

Figure 1: Youngest CEOs
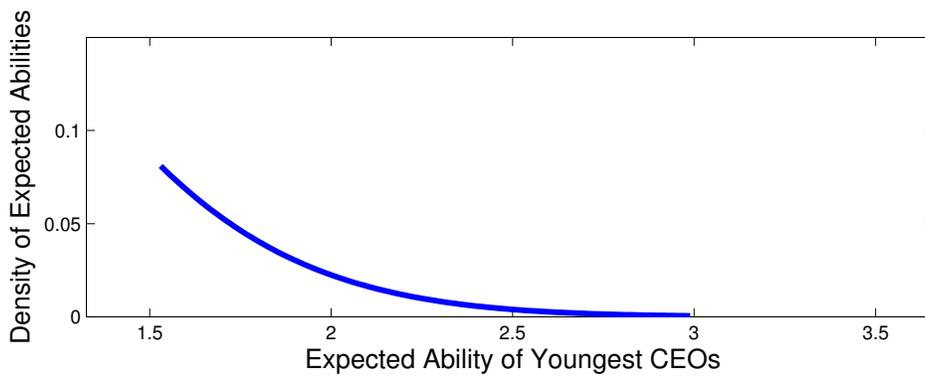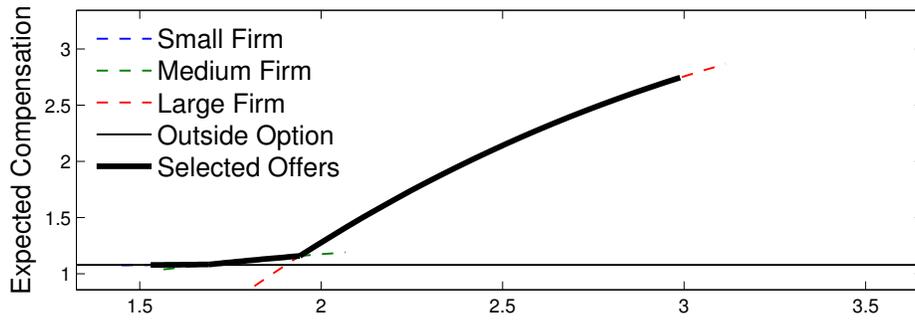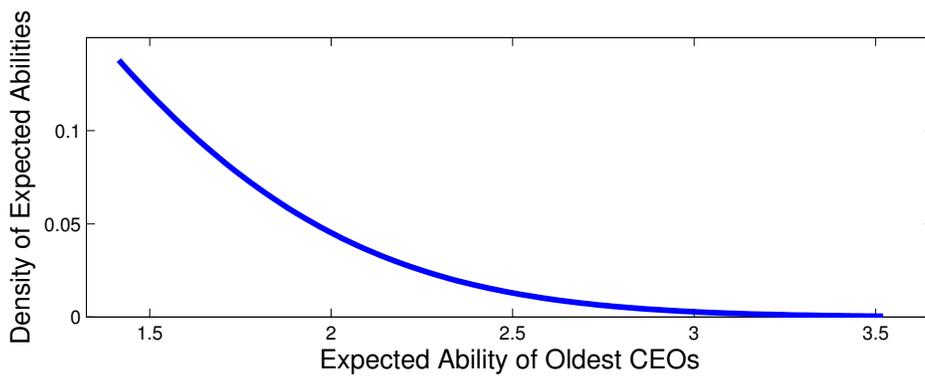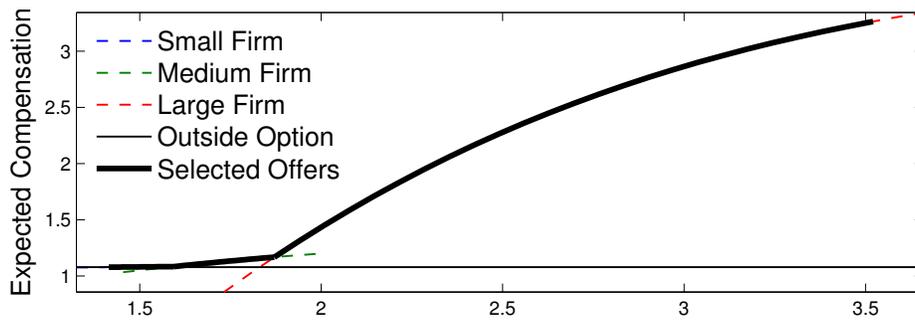

Figure 2: Oldest CEOs
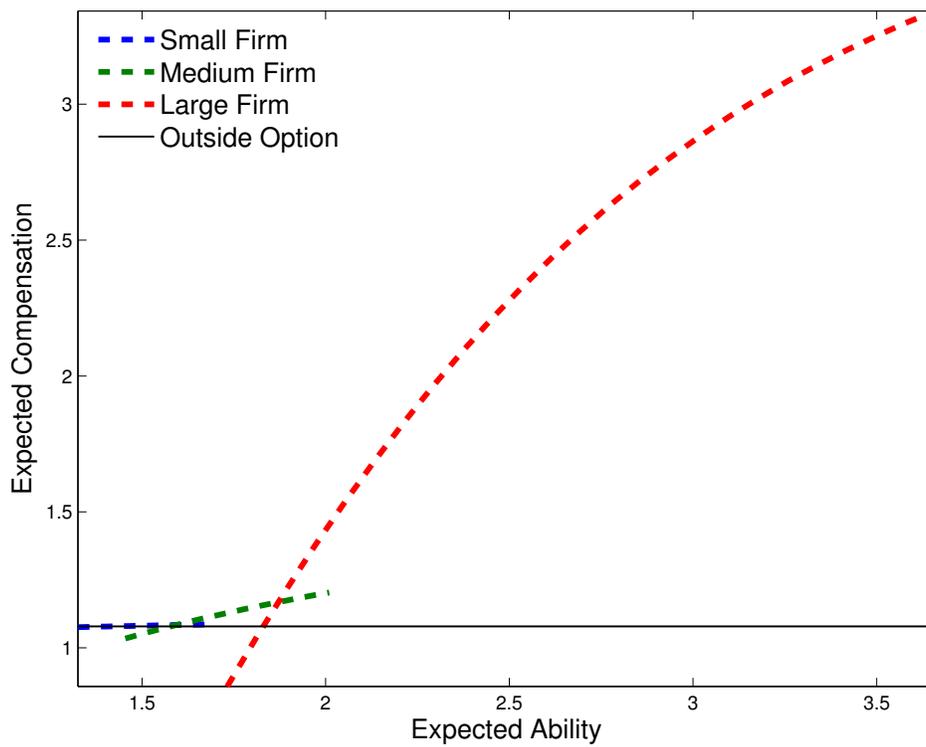
Figure 3: Youngest CEOs

Figure 4: Oldest CEOs

By combining (14) and (15) we can characterize the discrete increase in expected profits between firm $k-1$ and k by the difference in the expected match value when q_{k-1} and q_k are paired with the critical CEO indexed by $i_k(s)$:

$$\mathbb{E}\Big[x\Big(a_t\big(s,i_k(s)\big),q_k\Big)\Big] - \mathbb{E}\Big[x\Big(a_t\big(s,i_k(s)\big),q_{k-1}\Big)\Big]$$
$$= \mathbb{E}\Big[\pi_k\Big(a_t\big(s,i_k(s)\big)\Big)\Big] - \mathbb{E}\Big[\pi_{k-1}\Big(a_t\big(s,i_k(s)\big)\Big)\Big] \qquad (16)$$

The cross-cohort aggregation continues to follow (10) and the participation constraint is:

$$\mathbb{E}\big[x\big(a_t(s,i_1(s)),q_1\big)\big] = w + \mathbb{E}\Big[\pi_1\Big(a_t\big(s,i_1(s)\big)\Big)\Big] \qquad (17)$$

Since the market is long in marginal firm types, free entry guarantees that their expected profits are zero by the same argument as before.

Definition 2 (Equilibrium with Type-Contingency) *An equilibrium is a set of wage contracts $\{\omega_k(s,i)\}_{k=1}^K$ that satisfy the participation constraint (17) and the market clearing conditions in equations (12) and (13) by way of the indifference conditions in (15) and the cross-cohort aggregation in (10).*

Just as with output-contingent offers, CEOs select an offer that compensates them for their expected ability on the envelope. In contrast, since the contract payout does not depend on the realizations of the stochastic processes, the realized compensation is always a point on the envelope of offered contracts (see the plot of "Selected Offers" in figures 3 and 4 for young and old CEOs in our simple example).

Since everyone is risk-neutral, the type of contract on offer does not affect the optimal assignment of CEOs to firms.

Proposition 1 *The ex ante sorting of CEOs to firms is invariant to the identity of the residual claimant.*

Proof See appendix (forthcoming). Q.E.D.

3 Patterns of Mobility

The mobility patterns of CEOs in this model are governed by the evolution of the age-specific cross-sectional distribution of beliefs, the precision of the individual's point estimate, and the amount of curvature in the production technology $x(\cdot,\cdot)$. While matching is assortative in beliefs across members of a given age cohort, the dynamic aspects are much richer due to the

evolution of the distribution and precision of beliefs and their interactions with the projects' "appetite" for risk.

TO BE COMPLETED.

4 Data

We use information on CEOs and firm accounting data from three different data sources. The first is the Danish administrative register data covering 100% of the population of individuals and firms in the years 2000 to 2009. The second is from *Købmandstadens Oplysningsbureau* (KOB), which contains accounting data from the firm population dating back to 1991. The third is from the Danish Commerce and Companies Agency (*Erhvervs- og Selskabstyrelsen*, or ES) at the Ministry of Economic and Business Affairs and has information on managers of firms from 2000 to 2010. Personal information on managers can be linked from the ES data to the register data through a personal id number and the information on firm accounting data can be linked from the KOB data to the register data through a firm identifier. The register data further contains a link between the firm and the manager identifiers.

Using all three datasets to we can identify the CEOs of all Danish limited liability firms and link this to information on CEO compensations well as accounting data for the firms the CEOs work in.

Our first dataset is administrative register data, from Statistics Denmark, covering 100% of the population in the years 2000 to 2009. The Statistics Denmark data is from the Integrated Database for Labor Market Research (IDA), which contains annual information on socioeconomic variables (e.g., age, gender, education, etc.), characteristics of employment (e.g., wages, earnings, occupations, industries, etc.), and a employer-employee link for the population. The employer-employee link only exist for those individuals who held a job during the last week of November in any given year.

We use a measure of yearly earnings that contains yearly salaries including perks, tax-free salaries, jubilee- and severance payments, and the value of stock options and futures. These values are recorded when they are taxed, which happens at the year of creation and the year of exploitation (when they get sold or ceded/waived). The measure of earnings further contains payment for work on any board of directories - a payment we expect to be able to exclude for future work. Payments for consultancy work and other work related to giving presentation etc. are not included in the earnings variable.

The Statistics Denmark individual register data can be linked, through a firm identifier, to our second dataset, KOB, that contains firm accounting data, covering 100% of the limited liabil-

15

ity firm population in since 1991. The KOB dataset is collected and digitalized from scanned documents at the Ministry of Economic and Business Affairs by a private data collector called Experian. According to Danish law, all limited liability firms must provide the Ministry of Economic and Business Affairs with information about firms' asset and measures of profitability such as operating cost and net income. Other types of firm information (eg. number of employees, total wages, and sales) is provided voluntarily for a large majority of the firms. The firms' reported accounting data are subject to random auditing by external accountant which makes the reliability of the accounting data high.

Our third dataset, ES, is provided by the Danish Ministry of Economic and Business Affairs and contains information on all CEOs and board members in the limited liability firm in Denmark as well as information about the founders of the firms. The dataset contains information about the starting and finishing date of the spell for each CEO, board member, and founder and it includes an individual id number such that it is possible to link the the managers and founders, to their background characteristics in the Statistics Denmark dataset. The dataset also contains information on the type of firm the managers work in and the starting year of the firm.

We deflate the labor earnings, assets, and value added to the 2000 level using Statistics Denmark's consumer price index.

4.1   Sample Selection

We use privately and publicly held stock market firms in Denmark containing employees who work either part time or full time during the period 2000-2009. We then select those firms where the CEO has wage work in the firm as his or her main occupation during the last week of November of a given year.

While we can match a CEO to 94% of all stock market firms in Denmark, only 74% of the firms have a CEO who has his or her main employment in the firm. Since our focus is on the match between CEOs and firms, we exclude all firms without a CEO with a main employment in the firm. These excluded firms are primarily smaller than average since the included firms contains 80% of all the workers. Of the leftover firms, 8% has more than one registered CEO and for these firms, we identify the top CEO as the CEO with the highest occupational code (eg. "Directors and chief executives" is a higher occupational code than "Productions and operations manager")[8]. If a firm has two or more registered CEOs who have the same highest occupation code, we identify the CEO as the one who has the highest earnings. For 0.2% of the firms there are two or more CEOs with the same highest earnings and we also exclude

---

[8]The occupation code are from the Danish version of the ISCO codes.

these firms.

For our estimation we need accounting information from the firm and we therefore match the Statistics Denmark data to the accounting data from KOB. We match 95.5% of workers and 93.5% of the firms where the unmatched firms are smaller than average or firms that terminate during our selected sample period. We exclude the firms with unmatched KOB accounting data and we also exclude firms in the years they have missing information on value added, total assets, or fixed assets. There are less than 1% of firm-years with missing data on total assets or fixed assets and there are 4% of firm-years with missing value added. The firms with missing information are also primarily firms that are in their termination year or firms that are new in the sample. In total this gives a base sample of 149,729 firms that all have non-missing accounting data and an identified CEO whose main occupation is as CEO of the firm.

We think of a CEO as one who manages people and we therefore select CEOs from firms with five or more employees, which reduces the sample of firms to 115,608. In order to avoid complication with CEO retirement we further exclude all firms with CEOs of age 50 and above. Since we study CEO mobility between years, the sample only includes CEOs who also are CEOs in the data in the year after we use them in the analysis. This gives a final sample of 55,474 firms employing 2.8 million workers. Descriptive statistics of the CEO sample and the firm sample used in the analysis are provided in Table 1 and Table 2.

In order to match moments from the data, we divide the firms into bins by using deciles created either by the number of employees or values added.[9] The bins are computed using the firms in our sample and we use these bins when we calculate the moments relating to CEO transitions.

4.2   Empirical Moments

We use five empirical moment in our estimation, which are as follows. Roberts Law, which is the elasticity of CEO pay with respect to firm size, is calculated by using either employment or value added as firm size. We find that the elasticities are around 0.27 and are robust to conditioning on 1 to 4-digit industry fixed effects. The tail index of the Generalized Pareto distribution of firm sizes is similarly calculated using either employment or value added as our measure of firm size. In the former case, the parameter ξ is estimated at 0.63 while in the latter the index is one half.

The average CEO share of total output is 0.37 and is calculated by finding the average of CEO

---

[9]For the future we also use intangible assets.

	Mean	Median	25 percentile	75 percentile
Number of CEOs	55,474			
CEO to CEO switchers across firms	0.028			
CEO to CEO switchers across 10 employment bins	0.024			
CEO to CEO switchers across 10 value added bins	0.025			
Males	.94			
Age	41.6	42	38	46
Potential Experience	21.7	22	18	26
CEO tenure in firm	5.08	5	3	7
Hourly nominal wage in 2000 DKK	376.3	301.2	211.8	448.2
Total nominal earnings in 2000 DKK	623,270	499,751	347,884	746,664
Total nominal income in 2000 DKK	833,777	623,694	438,144	925,049

Table 1: Summary Statistics for CEOs

	Mean	Median	25 percentile	75 percentile
Number of firms	55,474			
Number of employees	50.6	18	10	37
Total assets in 1,000 DKK	195,397	11,406	5,471	28,319
Intangible assets in 1,000 DKK	157,312	7,153	3,306	18,276
Value added in 1,000 DKK	25,479	7,255	3,929	15,519

Table 2: Summary Statistics for Firms

earnings over the sum of CEO earning and the flow value of intangible assets. [10]

The annual wage growth of CEO compensation over the lifecycle is found by calculating each CEO's yearly earnings growth. We then either average across individuals in the sample of CEOs or we regress the earnings changes on calendar dummies and interpret the constant in the regression as the yearly earnings growth after controlling for calendar year effects. The simple average of earnings changes gives a yearly wage growth of 1.4% while controlling for calendar years boosts the rate to 1.7%.

We calculate the CEO-to-CEO transition probability by keeping track of switches across firm deciles, *conditional* on switching firms. For our purposes a switch is a CEO-to-CEO transition happens between year t and $t + 1$ or between t to $t + 2$, separated by a year of not being a

---

[10]If we use the value added as total output the CEO's compensation share is 0.141.

CEO in the sample. The reason for allowing a transition year is a peculiarity with respect to occupational classifications in our sample. CEOs are labeled as such if their job in November is their *main* employment during the year. By allowing a transition year we can account for CEO-to-CEO transitions that take place toward the of the year, but before the end of November. In this case, their November employment is not their main employment for that year. In the CEO-to-CEO transitions we therefore condition the CEO in year t on also being a CEO in year $t+1$ or $t+2$. We calculate the overall CEO-to-CEO transition probability across firm as 2.79% and CEO-to-CEO transition probability across firms *and* decile bins as 2.44%.

5 Calibration

In our calibration strategy we set five parameters in order to match five empirical moments and their model counterparts: the substitution elasticity ρ and the share parameter λ in the surplus technology, the Generalized Pareto tail index ξ of the distribution of firm sizes, the ratio of outside options $(\frac{p}{w})$, and the precision ψ of the normal innovations. The precision of true types is normalized to unity. The model is exactly identified and the five moments we aim to match are:

1. *Roberts' Law*, i.e. the elasticity of CEO pay with respect to firm size (employment or value added),

2. the tail index of the Generalized Pareto distribution of firm sizes (employment or value added),

3. the average CEO share of total output (match surplus),

4. the probability of CEO-to-CEO transitions, and

5. the average lifecycle growth rate of CEO compensation.

To compute individual spells we construct employment histories where we keep track of the assignments to deciles (vingtiles) in the firm size distribution rather than firm assignments themselves. In the model, CEOs have no special attachment to any firm. Firms are labels and thinking in terms of ranks along the distribution of firm sizes rather than firms is the natural empirical counterpart. This allows us to abstract from potential dynamics of the projects in the model. Unless ownership and management interact in ways that undermine frictionless separation and matching, this simplification has no impact on the structural estimates.[11] In

---

[11]This is an area for future research. A dynamic theory of entrepreneurship, for instance, requires that we think about the dynamics of productivities more carefully. This is part of an ongoing research agenda in collaboration with Moritz Meyer-ter-Vehn and Lee Ohanian.

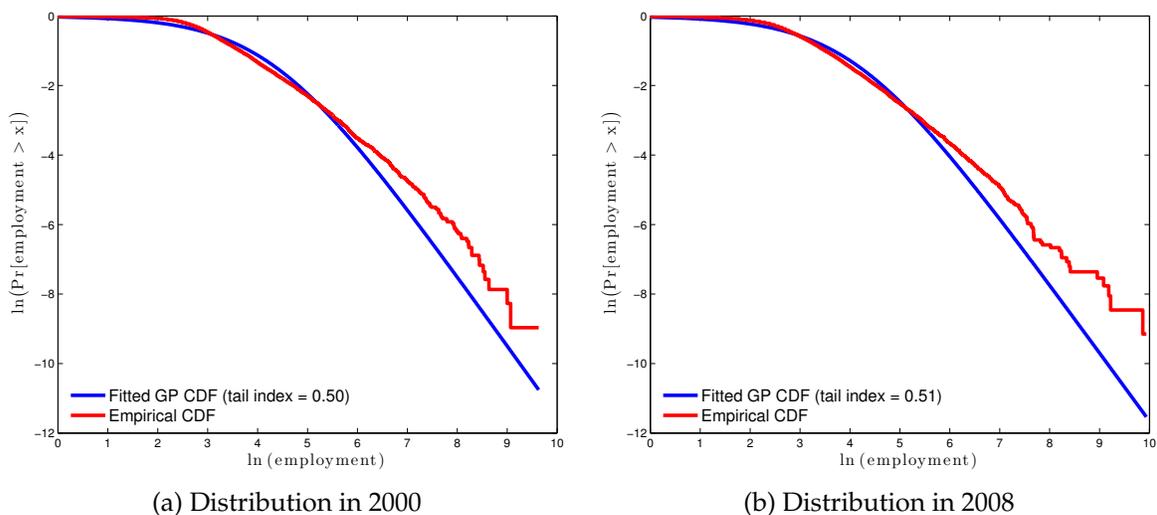(a) Distribution in 2000 (b) Distribution in 2008

Figure 5: Danish Firm Sizes

the data, spells are terminated when CEOs retire, when they switch from an employer in one decile (vingtile) to another employer in a different decile, or when their current employer moves to a higher or lower decile (vingtile) in the distribution of firm sizes without separation. Through the lens of our model, the spell of a CEO is not terminated when she moves from one firm to another in the *same* decile or vingtile and we classify turnovers in the data accordingly.

In our benchmark estimation, we use intangible capital to proxy for the firm size. We target an elasticity of 27% between the CEOs' compensation and the firms' intangibles . The point estimates hardly vary when we use employment instead. The average share of the CEOs' compensation relative to intangibles is 14%.[12] If instead, we use the firms' intangible assets as a proxy for q and compute the CEO's share relative to the annuity value of intangibles, the estimate more than doubles to 27%. We compute the maximum likelihood estimate of the Generalized Pareto (GP) tail index using employment rather than value added. We drop all those firms for which we cannot identify a CEO and our point estimates are 0.50 in 2000 and 0.51 in 2004 and 2008. We target a value of .50 in our benchmark. Figure 5 shows that the GP distribution is a good approximation of the size distribution of Danish firms with the usual caveat that it tends to underestimate the probability mass in the far right tail.[13]

TO BE COMPLETED.

---

[12]Estimates in Gabaix and Landier (2008) and Alder (2013) are around 30% for firms included in the Execu-Comp database of large US corporations. We have not computed the elasticity for a comparable Danish subsample.

[13]This is a standard observation in the literature on the distribution of firm and establishment sizes. If we include firms without known CEOs the estimate of the tail index rises to the 0.61-0.63 range.

6 Counterfactual Experiments

TO BE COMPLETED.

7 Conclusion

TO BE COMPLETED.

References

ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–333.

ALDER, S. D. (2013): "In the Wrong Hands: Complementarities, Resource Allocation, and Aggregate TFP," Working Paper.

BENNEDSEN, M., F. PÉREZ-GONZÁLEZ, AND D. WOLFENZON (2008): "Do CEOs Matter?" *Working Paper*, 1–43.

BENNEDSEN, M., F. PEREZ-GONZALEZ, AND D. WOLFENZON (2012): "Evaluating the Impact of The Boss: Evidence from CEO Hospitalization Events," Working Paper.

EECKHOUT, J. AND P. KIRCHER (2011): "Identifying Sorting – In Theory," *Review of Economic Studies*, 78, 872–906.

EISFELDT, A. L. AND C. M. KUHNEN (2013): "CEO Turnover in a Competitive Assignment Framework," *Journal of Financial Economics*, 109, 351–372.

EVANS, D. S. AND B. JOVANOVIC (1989): "An estimated model of entrepreneurial choice under liquidity constraints," *Journal of Political Economy*, 97, 808–827.

EVANS, D. S. AND L. S. LEIGHTON (1989): "Some empirical aspects of entrepreneurship," *American Economic Review*, 79, 519–535.

GABAIX, X. AND A. LANDIER (2008): "Why Has CEO Pay Increased So Much?" *Quarterly Journal of Economics*, 123, 49–100.

GROES, F., P. KIRCHER, AND I. MANOVSKII (2012): "U-shapes of Occupational Mobility," Working Paper.

JOHNSON, W. R. (1978): "A theory of job shopping," *Quarterly Journal of Economics*, 92, 261–278.

JOVANOVIC, B. (1979): "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 87, 972–990.

KNIGHT, F. H. (1921): *Risk, Uncertainty, and Profit*, New York: Houghton Mifflin.

LISE, J., C. MEGHIR, AND J.-M. ROBIN (2013): "Mismatch, Sorting and Wage Dynamics," *Working Paper*.

LOPES DE MELO, R. (2013): "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence," *Working Paper*, 1–39.

LUCAS, R. E. (1978): "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9, 508–523.

MILLER, R. A. (1984): "Job matching and occupational choice," *Journal of Political Economy*, 92, 1086–1120.

NAGYPAL, E. (2007): "Learning by doing vs. learning about match quality: Can we tell them apart?" *Review of Economic Studies*, 74, 537–566.

RUGE-MURCIA, F. J. (2010): "Estimating Nonlinear DSGE Models by the Simulated Method of Moments," *Working Paper*, 1–59.

TERVIÖ, M. (2008): "The Difference That CEOs Make: An Assignment Model Approach," *American Economic Review*, 98, 642–68.