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

Simeon D. Alder*  
University of Notre Dame

Fane N. Groes†  
Copenhagen Business School

February 28, 2013

PRELIMINARY AND INCOMPLETE – PLEASE DO NOT CITE

Abstract

We build a tractable assignment model to identify the degree of complementarity between attributes of firms and managers in production. 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 use a simulated method of moments to replicate empirical targets from a large matched employer-employee dataset covering the Danish labor force between 1999 and 2007. We exploit the non-monotonicity of executive compensation in the employer type to structurally estimate the degrees of complementarity in the production function. Our results fill an important gap in the literature on the aggregate effects of mismatch.

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 (heterogeneous) skills across (heterogeneous) firms and jobs. Lucas (1978)’ model on the allocation of workers and capital to a manager is a canonical explanation for the distribution of firm sizes. It emphasizes the intensive capital and labor margin and has given rise to quantitative work on the effects of factor misallocation on aggregate productivity (see, for instance, Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009). Recent work by Tervio (2008) and Gabaix and Landier (2008) builds on Becker (1973) and Sattinger (1979). In this literature firm size is determined by the joint contribution of the manager’s and the firm’s attributes. In the presence of two-sided heterogeneity, the distribution of types and the extent to which complementarities govern the value created by any manager-firm pair are key to understanding sorting and compensation patterns. If managers and firms are indeed complements, the compensation of CEOs and shareholders depends on the attributes of both. By the same token, individual (or marginal) effects are no longer identified easily and that poses a challenge. To the extent that we are concerned with efficiency, complementarities are crucial. If the technology that governed the joint value of a manager and his job was linear, a CEOs marginal product would simply be a function of the efficiency units she embodies. Her contribution to value is the same across employers and the assignment of managerial talent to jobs is of no consequence for efficiency. In such a world, Guy Laliberté the street performer has the same value as Guy Laliberté the CEO of Cirque du Soleil. However, we have reason to believe that complementarities matter. While Steve Jobs’ input made a difference at Pixar, he had an even more significant impact at Apple.

Assignment models in the tradition of Becker (1973) and Sattinger (1979) typically abstract from the intensive margin: managers run firms without capital or labor inputs. The presence of both the intensive and extensive margins, however, enables us to address salient macroeconomic questions. What are the implications for efficiency when someone like Guy Laliberté is stuck with a location (the streets of Quebec City) that can support a one-man show rather than the state-of-the-art stages and tents that allow Cirque du Soleil to entertain audiences with scores of artists, craftsmen, and musicians? What effect, if any, do frictions or policies have that stifle Guy’s entrepreneurial spirit or force him to enchant his target audience(s) one at a time rather than with simultaneous performances around the world? Alder (2012) and Eeckhout
and Kircher (2012), among others, combine the intensive and extensive margins in a single model and show how complementarities are key to understanding the effects of mismatch and misallocation on the distribution of firm sizes and on macroeconomic productivity. In these models, the challenge is to identify the distribution of types on both sides of the market separately from the technology that determines the value of any firm-worker pair. This is well-documented feature of assignment models with equilibria that exhibit perfect assortative matching and this paper proposes a model that can overcome this limitation.

Our mechanism has potential applications beyond the CEO-to-firm assignment problem considered here. We can account for the stylized effects of skill-biased technological change in terms of hiring patterns and wage inequality, for instance, in a framework that is different, but complementary, to Krusell et al. (2000). In this regard, our work is in the same spirit as Eeckhout and Kircher (2012).

We set up an assignment model where agents do not know their own type but receive a sequence of public signals that enables them to Bayesian update their beliefs (as in Groes et al., 2012). While the equilibrium features perfect assortative matching ex ante, it does not do so ex post, once a new public signal has been observed, that is. Re-sorting is friction- and costless. Each period, the agents can pick the employer that offers the most lucrative wage contract, in expectation. As in Eeckhout and Kircher (2011), compensation is non-monotonic in the employer type and we explore this property to characterize (re-)sorting patterns and to estimate the degree of complementarity with comprehensive panel data for CEOs and their employers. We use a simulated method of moments based on a large matched employer-employee dataset covering the Danish labor force from 1999 to 2007.1 Much of the existing literature ignores the firm data on output, profits, employment, or value added since it does not respond much to variations in the productivity of individual jobs. Since we consider the matching of a worker to a unique job in the firm, the chief executive officer, we can not only characterize the strength of sorting, but also the sign (positive vs. negative assortative matching).

The remainder of the paper is organized as follows. Section 2 details the model environment and defines the equilibrium. We describe the data in section 3 and the

---

1We are in the process of adding data that extends the range for some our variables as far back as 1980. This does not affect our theory but requires us to update sections 3 and 4 in due time.
specifics of our structural estimation in section 4. Section 5 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 0 and precision $\phi$. The draws from this distribution are i.i.d. across agents. In addition, they have a periodic unit endowment of raw labor.

Individuals live for $T$ periods. Each period, a new cohort of size $\frac{1}{T}$ enters the labor market, 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 true type $a$ follows $N(0, \phi)$ and the noise $\alpha$ is normal with mean zero and precision $\psi_0$. At time $t \in \{1, \ldots, T\}$, the managers’ contribution to a match surplus – if any – is $a_t = a + \alpha_t$. The uncertainty associated with the CEOs’ input, $\alpha_t$, is drawn from a normal distribution with mean zero and precision $\psi$. Each period, individuals decide whether to supply a raw labor unit or their managerial talent.\textsuperscript{2}

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 bequested to randomly chosen individuals of the cohort with birth date $s + T$.

\textsuperscript{2}Nothing prevents them from splitting their time between wage labor and managerial work. In equilibrium, however, their occupational choice is binary. In future work, we plan to explore the implications of a model where individuals are endowed with an $N$-dimensional vector of abilities. Each $a_m$ denotes the ability to carry out one of the $N$ executive jobs in a firm. The surplus (or output) generated by a team of managers $(a_1, \ldots, a_N)$ in charge of a project $q$ depends on everyone’s contribution. Rather than equation (1), the production function of this model is given by

$$x_t = \left[ \lambda f\left(\{a_i, t\}_{i=1}^{N}\right)^{\frac{\alpha - 1}{\sigma}} + (1 - \lambda)(q + \theta_t)^{\frac{\alpha - 1}{\sigma}} \right]^{\frac{1}{\alpha - 1}}$$

with $f\left(\{a_i, t\}_{i=1}^{N}\right) = \left( \sum_{i=1}^{N} \eta_i (a_i + \alpha_{i,t})^{\frac{\alpha - 1}{\sigma}} \right)^{\frac{1}{\alpha - 1}}$ and $\sum_{i=1}^{N} \eta_i = 1$. This is, in essence, a Roy model with non-linear prices (executive pay schedules).
2.2 Preferences and Technology

Individuals have linear preferences over their lifetime consumption stream:

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

Since there is no consumption-saving tradeoff, \( c_s \) is simply capped by the sum of labor and executive income as well as dividend payments and capital gains from project ownership. A CEO-project match produces the consumption good subject to a pair of shocks \((\alpha_t, \theta_t)\):

\[ x(a + \alpha_t, q + \theta_t) = \begin{cases} 
\left( \lambda(a + \alpha_t)^\frac{\psi-1}{\rho} + (1 - \lambda)(q + \theta_t)^\frac{\psi-1}{\rho} \right)^{\frac{\rho}{\psi-1}} & \text{if } a + \alpha_t \geq 0, q + \theta_t \geq 0 \\
0 & \text{otherwise.} \end{cases} \]  

(1)

Shareholders and CEOs have limited liability. In particular, a firm’s shareholders cannot appropriate income generated by another project, owned by the firm’s CEO but run by another executive.

2.3 Beliefs

The distribution of managerial talent is stationary. Each cohort of size \( \frac{1}{T} \) draws from the full support of a normal distribution with mean zero and variance \( \frac{1}{\psi} \). Clearly, 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. To keep the model simple, we assume that the signal at birth \((t = 0)\) and those later in life follow the same process with precision \( \psi \):^3

\[ \alpha_t \sim N(0, \frac{1}{\psi}), \text{ for all } t \]  

(2)

Each signal enables individuals to update their prior using a Kalman filter. Since the innovation is normal, the posterior of an age-\( s \) – denoted by \( \tilde{a}_{s+1} \) – is normal with

---

^3Relaxing this assumption generates the same qualitative results but requires heavier notation.
mean
\[ \hat{a}(s + 1) \equiv \mathbb{E}(\hat{a}(s + 1)) = \frac{\phi_i + s\psi}{\phi_i + (s + 1)\psi} \hat{a}(s) + \frac{\psi}{\phi_i + (s + 1)\psi} a_t \] (3)

and variance
\[ \text{Var}(\hat{a}(s + 1)) = \frac{1}{\phi + (s + 1)\psi} \] (4)

Note that the variance of the posterior is 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.\(^4\) 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 age.

## 2.4 Compensation Offers and Firm Profits

Since the model is universally risk-neutral matching decisions are based on expected payments to managers and project owners. That insight does not, however, identify the residual claimant \textit{ex post} and we can choose between two polar cases:

1. Wage offers are output-contingent contracts. The firm holds on to a reservation profit and the CEO is the residual claimant.

2. Wage offers are \textit{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 \textit{ex post} surplus require that project owners know as much about the applicants as the CEOs themselves do. If they do not, CEOs have incentives to misrepresent their type in order to extract more surplus.

\(^4\)The derivations of the distribution of individual posteriors as well as the cross-sectional distribution of priors (which will be discussed shortly) are available from the authors upon request.
Clearly, the realized split of the surplus between managerial compensation and payments to shareholders depend on the type of contract. In particular, the predicted patterns of pay volatility and switching differ between the two types of contracts. We spell out the wage offers and profits under self-selection and full insurance and let the data point us in the right direction. In the presence of information frictions and risk-aversion, one can imagine that firms may want to provide incentives for self-selection in combination with some insurance against idiosyncratic risk. In this case, the wage offer may be a convex combination of the two polar contracts.\(^5\)

Before we spell out the two types of offers, it is useful to go over some preliminaries and introduce additional notation.

Beliefs about a CEO’s type are denoted by \(\tilde{a}(s)\) 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 \(T\). In the special case of our model where the production function is linear (\(\rho \to \infty\)), we can ignore the variance altogether. When the technology is (strictly) super-modular, however, we need to keep track of a CEO’s expectation and variance. Since the innovation \(\alpha_t\) is normal, the prior \(\tilde{a}\) is also normal.\(^6\)

\[
\tilde{a}(s) \sim N\left(\hat{a}, \frac{1}{\phi + s\psi}\right) \tag{5}
\]

The expected output of a manager with prior \(\tilde{a}(s)\) who is paired with a project \(q\) is denoted by

\[
\mathbb{E}\left[x(\tilde{a}(s), q)\right] \equiv \int_{-q}^{\infty} \int_{0}^{\infty} \left[\lambda \tilde{a}(s)^{\frac{\rho - 1}{\rho}} + (1 - \lambda) (q + \theta)^{\frac{\rho - 1}{\rho}}\right]^{\frac{\rho}{\rho - 1}} dF_{\tilde{a}(s)}(a) dG(\theta) \tag{6}
\]

where \(F_{\tilde{a}(s)}\) is the C.D.F. of \(\tilde{a}\) and \(G\) is the normal distribution function of \(\theta\) with mean zero and precision \(\psi\).\(^7\)

---

\(^5\)Here, we ignore asymmetric information and risk-aversion. The model, therefore, does not point us in the direction of self-selection or insurance. For the time being, we simply focus on the two polar contracts and ignore the possibility of intermediate cases.

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

Under output-contingent offers, CEOs are residual claimants and bear all the risk. Firms keep the reservation profit \( \pi(q) \) and offer \( \omega(x) = x(a + \alpha_t, q + \theta_t) - \pi(q) \) to prospective CEOs. Keep in mind that for particularly unfavorable realizations of \( a + \alpha_t \) and \( \theta_t \), 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 such that the vacancy is filled. That is, the \( K \) different \( \pi(q) \) 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:

\[
\hat{a}(s) \sim N(0, \frac{s\psi}{\phi(\phi + s\psi)})
\]

We denote the corresponding C.D.F. by \( F_{\hat{a}(s)} \). Recall that the true (but unobserved) distribution of abilities is \( N(0, \phi^{-1}) \). The cross-section of beliefs must be unbiased and 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 expected abilities converge to the true types \( a \) with a variance that decays to zero as \( s \to \infty \). In the limit, CEOs know their own type for sure. The distribution of expected abilities has variance \( \frac{s\psi}{\phi(\phi + s\psi)} \) and converges to the cross-section of true types. In particular, the variance of beliefs approaches the variance of true types:
Since the sorting is \textit{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 CDF of expected abilities in cohort \( s \).

\[ \hat{a}(s)[i] = F^{-1}_{\hat{a}(s)}(i) \quad (8) \]

Let \( j_k(s) \) index a CEO in cohort \( s \) and in a slight abuse of notation we denote the expected output of a CEO with expected ability \( \hat{a} \) by \( \mathbb{E}[x(\hat{a}(s)[j_k(s)], q_{k-1})] \). For a given \( \{\pi(q_k)\}_{k=1}^{K} \), this critical CEO is indifferent between managing the project \( q_{k-1} \) and the “adjacent” project \( q_k \). Formally, there is a set \( \{j_k(s)\}_{k=1}^{K} \) that satisfies:

\[ \mathbb{E}[x(\hat{a}(s)[j_k(s)], q_{k-1})] - \pi(q_{k-1}) = \mathbb{E}[x(\hat{a}(s)[j_k(s)], q_k)] - \pi(q_k) \quad (9) \]

Let \( j_0(s) = 0 \) for all \( s \). Then \( \frac{j_k(s) - j_{k-1}(s)}{T} \) denotes the measure of CEOs who select project \( k \). Given the beliefs about their own type, alternative employers offer less lucrative contracts in expectation. Summing over all cohorts delivers the total measure of those who assign themselves to project \( k \):

\[ j_k - j_{k-1} = \frac{1}{T} \sum_{s=1}^{T} j_k(s) - j_{k-1}(s) \quad (10) \]

In addition, we impose a participation constraint. CEOs have an occupational outside option: they can supply their raw labor unit and earn the competitive wage \( w \). Projects have the outside option to idle and generate zero profits. Clearly, then, the marginal CEOs indexed by \( j^*_k(s) \) and the marginal project \( k^* \) satisfy:

\[ \mathbb{E}[x(\hat{a}(s)[j^*_k(s)], q_{k^*})] = w \quad (11) \]

Knowing the distribution of beliefs across all cohorts, firms choose the reservation
profits \( \{\pi(q_\ell)\}_{\ell=k^*+1}^{K} \) that satisfy (11) and clear the market for CEOs:

\[
\begin{align*}
    j_{k^*+1} - j_{k^*} & \leq \gamma_{k^*} \quad (12) \\
    j_{\ell+1} - j_{\ell} & = \gamma_k, \text{ for } \ell = k^* + 1, \ldots, K - 1 \text{ and } j_{K+1} = 1 \quad (13)
\end{align*}
\]

**Definition 1 (Equilibrium with Output-Contingency)** An equilibrium is a set of reservation profits \( \{\pi(q_\ell)\}_{\ell=k^*}^{K} \) 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 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 \( K -(k^* - 1) = 3, T = 5 \) and outside option \( w = 0.986 \). This competitive wage implies that \( j_{k^*} = 0.95 \) and only the right tail of individuals are 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 (as in Groes et al., 2012, for instance).

Figures 3 and 4 highlight the ex ante selection: CEOs never accept a job that lies inside the envelope covering the \( K -(k^* - 1) \) 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. In the special case with \( \theta_t = 0 \) for all \( t \), the plot of realized ability \( a + \alpha_t \) against actual compensation lies on the wage offer curve, but not necessarily on the envelope. When \( \theta \) is non-degenerate, on the other hand, the plot of \( a + \alpha_t \) against \( \omega(x) \) is clustered around the envelope (both above and below).
Figure 1: Youngest CEOs

Figure 2: Oldest CEOs
Figure 3: Youngest CEOs

Figure 4: Oldest CEOs
2.4.2 Type-Contingent Offers

In this environment, a type is a distribution characterized by the pair \((\hat{a}, s)\), that is, a belief (expected ability) and an age (variance). When the offers are type-contingent, the project owners are residual claimants. Instead of picking \(\{\pi(q_{\ell})\}_{\ell=k^*}^{K}\), they select a set of expected profits \(\{E[\pi(q_{\ell})]\}_{\ell=k^*}^{K}\) that clear the market for CEOs ex ante. The CEOs are compensated according to their expected ability:

\[
\omega(\hat{a}(s)) = E\left[x(\hat{a}(s)[j_k(s)], q_k) - \pi(q_k)\right]
\]

(14)

The expected reservation profits must satisfy the critical CEOs’ indifference condition:

\[
E\left[x(\hat{a}(s)[j_k(s)], q_{k-1}) - \pi(q_{k-1})\right] = E\left[x(\hat{a}(s)[j_k(s)], q_k) - \pi(q_k)\right]
\]

(15)

In contrast to (9), the expectation is over output and profits. The cross-cohort aggregation continues to follow (10) and the participation constraint remains as in (11).

**Definition 2 (Equilibrium with Type-Contingency)** An equilibrium is a set of expected reservation profits \(\{E[\pi(q_{\ell})]\}_{\ell=k^*}^{K}\) that satisfy the participation constraint in (11) and the market clearing conditions in equations (12) and (13) by way of the indifference conditions in (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 projects, as long as all firms offer the same type of contract.

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

**Proof** To be completed.

Q.E.D.
The equilibrium has a number of additional properties. We simply claim them here and provide proofs later.

**Claim 1** Additional properties that both equilibria have in common:

- For a given innovation, older CEOs are less likely to switch to a different employer.

- Learning about CEO types implies mobility in both directions with respect to the firm size distribution. With strictly super-modular technology, CEOs are more likely to switch to a bigger firm, on average. In contrast, they are equally likely to switch in either direction when the technology is linear.

- Since agents are risk-neutral and learning is independent of the match, the dynamic equilibrium collapses to a sequence of static equilibria.

In our quantitative exercise we parameterize the equilibrium with output-contingent contracts to match moments in a matched employer-employee dataset. One of the reasons to select this equilibrium is that the data exhibits the kind of overlap in realized CEO pay that is consistent with such contracts in the model. In future work, we plan to verify if the results with respect to our key parameter of interest – the substitution elasticity $\rho$ – are robust to the alternative type-contingent contract.

## 3 Data

We use the administrative Danish register data covering 100% of the population in the years 1999 to 2007. The first part of the 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, occupations, industries, etc.), and a employer-employee link for the population. Wages consists of the hourly wage in the job held during the last week of November of each year. Wage information is not available for workers who are not employed in the last week of November. Statistics Denmark attached a measure of quality to the wages calculated from hours worked and information about general wages in the Danish labor market. Our results on wages include only employees with high defined quality of their wage information.
The firms linked to the workers consist of stock companies. One firm can consist of many workplaces. We use data on the firms from IDA’s workplace register and the register of firm statistics. The IDA workplace register contains information about ownership structure and ID for the workplace that it constant over time. The register of firm statistics is available for the period 1999-2007 and includes accounting data from the firm, such as total revenue and profit.

Occupational affiliation is defined by the so-called DISCO code, which is the Danish version of the ISCO-88 classification (International Standard Classification of Occupations). The validity of the codes is considered to be high, in particular, because they are monitored by employers and unions and form the basis of wage bargaining at the national level. We use the most disaggregated definition of the occupational classification available for the entire period, i.e., the 4-digit code. This classification corresponds fairly closely to the overall 3-digit Standard Occupational Classification used by the U.S. Census.

As a second way to classify CEOs, we use “primary work position”, which contains information about the level of employment. This variable includes seven different levels of employment from director of company to unskilled workers and non-classified other workers.

3.1 Sample Selection

Individuals are selected for the analysis if their primary work is as a wage worker during the last week of November in a given year and if they are between 18 to 70 years old.

We use all stock market firm in Denmark containing employees selected through the above criteria for the period 1999-2007. The firm register contains information on number of employees during the last week of November, which we use for firm size distributions. A firm can thereby have fewer observed/selected employees than number of employees.

A CEO is defined as the worker within the firm with the highest code in “primary work position”. For firms with two or more workers with the same highest code, the worker with the highest hourly salary is chosen to be the CEO. For about 5% of the

---

firms there are two or more workers with the same highest work position code and the same hourly wages. To define a CEO for these firms, we choose the one with the longest labor market experience. For this analysis, we have chosen only to include CEOs with a full time job at the firm but we keep all firms in the sample whether they have a CEO or not.

Descriptive statistics of the CEO sample and the firm sample used in the analysis are provided in Table 1 and Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>25 percentile</th>
<th>75 percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of CEOs</td>
<td>65,026</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CEO to CEO switchers</td>
<td>0.021</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>.90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years of school</td>
<td>14.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>46.1</td>
<td>46</td>
<td>39</td>
<td>54</td>
</tr>
<tr>
<td>CEO tenure</td>
<td>8.37</td>
<td>6</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>Hourly nominal wage in DKK</td>
<td>418.6</td>
<td>336</td>
<td>245</td>
<td>491</td>
</tr>
<tr>
<td>Total nominal earnings in DKK</td>
<td>705,077</td>
<td>563,955</td>
<td>407,183</td>
<td>825,926</td>
</tr>
<tr>
<td>Total nominal income in DKK</td>
<td>763,152</td>
<td>589,556</td>
<td>425,103</td>
<td>867,006</td>
</tr>
</tbody>
</table>

Table 1: Summary Statistics for CEOs

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>25 percentile</th>
<th>75 percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of firms</td>
<td>99,565</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of employees</td>
<td>72.9</td>
<td>22</td>
<td>10</td>
<td>49</td>
</tr>
<tr>
<td>Gross profits in 1,000 DKK</td>
<td>48,862</td>
<td>12,258</td>
<td>5,754</td>
<td>29,289</td>
</tr>
</tbody>
</table>

Table 2: Summary Statistics for Firms
4 Estimation

We use a simulated method of moments (SMM) to structurally estimate the model. This quantitative work is currently in progress and we expect to report results shortly. The parameter that is of most interest is $\rho$, the elasticity of substitution between the CEO’s and the firm’s types.

Given the size of the parameter space we rely on a simulated annealing algorithm to minimize the curse of dimensionality. We target five moments in the data with corresponding (simulated) counterparts in the model:

1. Roberts’ Law: elasticity of CEO pay with respect to firm size (employment or value added)
2. tail index of the Pareto distribution of firm sizes (employment or value added)
3. average CEO share of total output (surplus)
4. average length of CEO spells in a firm size bin (or, equivalently, the probability of switching employers across firm size bins)
5. non-monotonicity of CEO compensation across employers in different size bins (measured by across-bin variations in Roberts’ Law)

Our preliminary estimates for Roberts’ Law are between 0.2 and 0.24, which is somewhat lower than the corresponding elasticity for large US corporations.\(^9\) Figure 5 shows that the size distribution of firms in Denmark has the familiar US shape – except for the handful of observations in the far right tail – and we use a (discrete) Zeta approximation to the Pareto distribution to generate the firm types in the model.\(^10\)

5 Conclusion

TO BE COMPLETED.

---

\(^9\)Estimates in Gabaix and Landier (2008) and Alder (2012) are around 30% for firms included in the ExecuComp database of large US corporations.

\(^10\)The matched employer-employee dataset is currently being updated by Statistics Denmark. We will be able to compute the remaining moments and initiate the structural estimation once the update is complete.
Figure 5: Danish Firm Size Distribution

References


